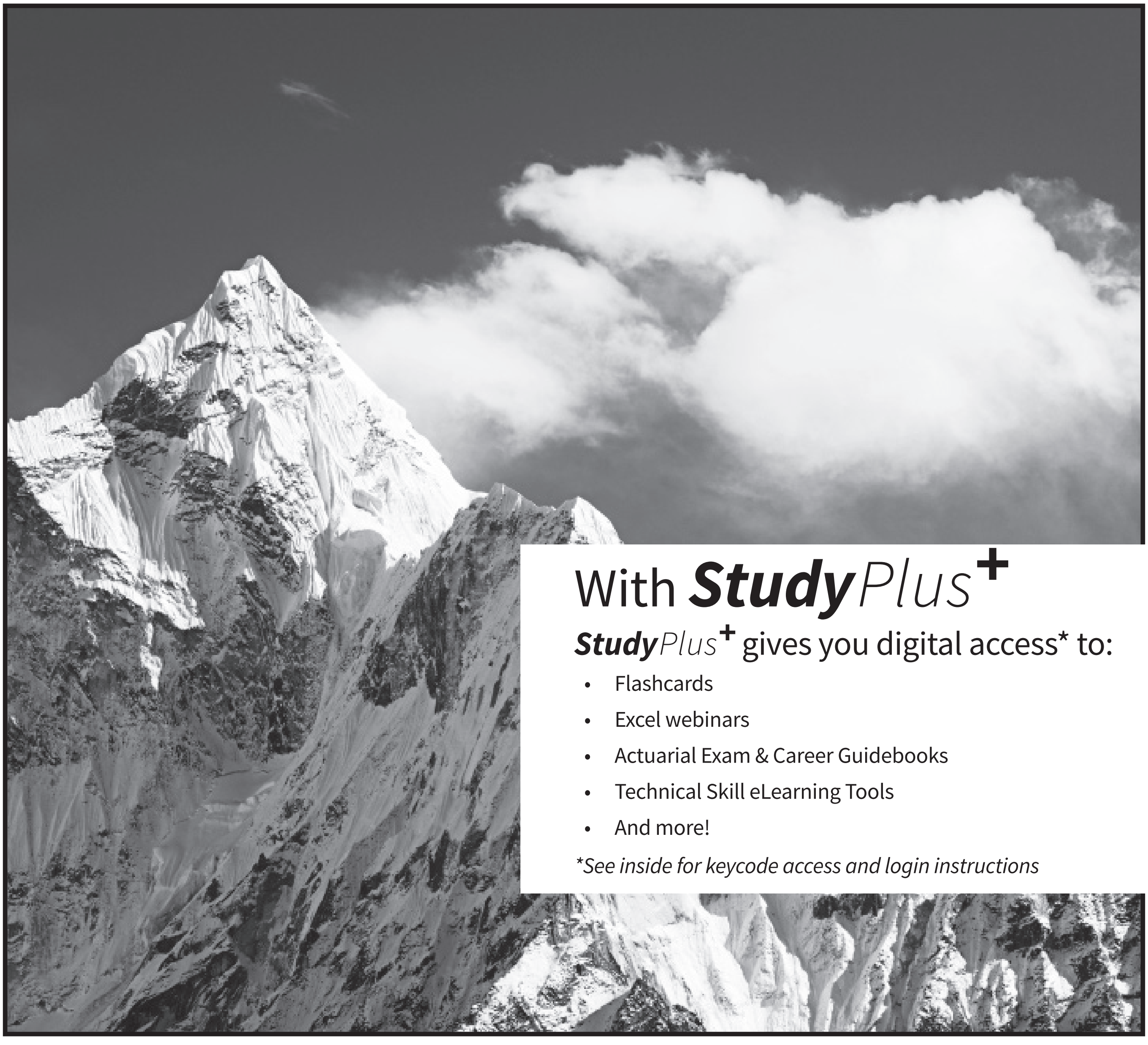
ACTEX

CAS Exam MAS-I Study Manual



With **Study**Plus+

StudyPlus+ gives you digital access* to:

- Flashcards
- Excel webinars
- Actuarial Exam & Career Guidebooks
- Technical Skill eLearning Tools
- And more!

**See inside for keycode access and login instructions*

Fall 2018 Edition | **Volume I**

Ambrose Lo, FSA, Ph.D., CERA

ACTEX Learning | Learn Today. Lead Tomorrow.

ACTEX Learning
Learn Today. Lead Tomorrow.

Actuarial & Financial Risk Resource Materials
Since 1972

ISBN: 978-1-63588-536-1

Printed in the United States of America.

Contents

Preface

Exam MAS-I (Modern Actuarial Statistics I) is a new exam which was offered for the first time in Spring 2018 by the Casualty Actuarial Society (CAS). In Fall 2018, the exam is scheduled for October 25, 2018. The registration deadline is August 30. This exam replaces its predecessor Exam S (*Statistics and Probabilistic Models*), which is a relatively short-lived exam offered only five times, from Fall 2015 to Fall 2017. Exam S, in turn, was developed from the old Exam LC (*Models for Life Contingencies*), Exam ST (*Models for Stochastic Processes and Statistics*), and Exam 3L (*Life Contingencies and Statistics*). The introduction of Exam MAS-I is in response to the discontinuation of Exam C/4 in July 2018, which the CAS sees as an opportunity to revamp Exam S with a heavier focus on contemporary statistical methods and the addition of statistical learning as a way to enhance the statistical literacy of property and casualty actuaries in this day and age. You will considerably sharpen your statistics toolkit as a result of taking (and, in all likelihood, passing!) Exam MAS-I.

Syllabus

The syllabus of Exam MAS-I, available from `http://www.casact.org/admissions/syllabus/ExamMASI.pdf`, is extremely broad (but not necessarily deep) in scope, covering miscellaneous topics in applied probability, mathematical statistics, statistical modeling and time series analysis, many of which are new topics not tested in any SOA/CAS past exams. As a rough estimate, you need at least *three months* of intensive study to master the material in this exam.[i] The specific sections of the syllabus along with their approximate weights in the exam are shown below:

Section	**Range of Weight**
A. Probability Models (Stochastic Processes & Survival Models)	20–35%
B. Statistics	15–30%
C. Extended Linear Models	30–50%
D. Time Series with Constant Variance	10–20%

Compared with the former Exam S, both Sections C and D have enjoyed a heavier weight, from 25–40% to 30–50% and from 5–10% to 10–20%, respectively. Sections A, B, and D are more or less taken intact from the syllabuses of Exam S, ST, LC, and 3L. As a result, you can

[i]It is true that one need not master every topic in order to pass this exam.

find lots of relevant past exam questions on these two sections. Section C has experienced the biggest change, with a new textbook on statistical learning added.

Exam Format

Exam MAS-I is a four-hour multiple-choice exam. According to the list of MAS-I frequently asked questions (`http://www.casact.org/cms/files/New_CAS_Exams_MAS_I_and_II_FAQs_1.pdf`), the exam will consist of approximately 35 to 40 questions. However, the Spring 2018 exam has 45 questions and you can expect future exams to have 45 questions as well. Before the start of the exam, there will be a fifteen-minute reading period in which you can silently read the questions and check the exam booklet for missing or defective pages. However, writing will not be permitted during this time, neither will the use of calculators.

Given the similarity between Exam MAS-I and Exam S (the syllabus of the latter is available from `http://www.casact.org/admissions/syllabus/ExamS.pdf`) in terms of their structure and topics, we may use Exam S as a rough proxy for Exam MAS-I. Each of the Exam S papers from Fall 2015 to Fall 2017 has 45 questions, categorized into the four sections as follows:

	Number of Questions					
Section	15F	16S	16F	17S	17F	18S
A. Probability Models	13	12	16	15	16	14
B. Statistics	17	15	14	13	14	9
C. Extended Linear Models	11	15	11	14	12	17
D. Time Series with Constant Variance	4	3	4	3	3	5
Total	45	45	45	45	45	45

You can see that roughly the same number of exam questions was set on Sections A and C, although Section C is proclaimed to be the most important section (perhaps even the examiners found it hard to set questions on this section!). To investigate whether such a distribution of exam questions is consistent with the distribution that the CAS announced in the exam syllabus, please try Practice Exam 2 Question 22 on page 1359. According to `http://www.casact.org/admissions/passmarks/examS.pdf`, the pass marks for Fall 2017, Spring 2017, Fall 2016, Spring 2016, and Fall 2015[ii] were **50.0**, **50.5**, **54.0**, **55.0**, and **52.5**, respectively, which means that candidates needed to answer about **26 to 27 out of 45 questions** correctly to earn a pass (each question carries 2 points with the total score being $44 \times 2 = 88$ or $45 \times 2 = 90$).

Here are the characteristics of a typical CAS multiple-choice exam:

1. The questions are almost always arranged in the same order as the topics in the exam syllabus, so Question 1 is very likely a Poisson process question and Question 45 is almost always a time series question. This implicitly gives you a hint as to which topic an exam question is testing.

[ii] The pass mark for the Spring 2018 Exam MAS-I is not yet available when this manual goes in press.

2. A number of exam questions bear a striking resemblance to past CAS exam questions, sometimes even with the same numerical values. For example, Questions 15 and 20 of the Spring 2018 MAS-I exam were direct modifications of Questions 18 and 21 of the Fall 2010 Exam 3L, respectively; only the numerical values differ. This attests to the importance of practicing numerous past exam problems, an abundance of which are carefully discussed and solved in this study manual.

3. The scope of an exam can be narrow at times with several questions testing the same topic in much the same way. For example, Questions 43, 44, and 45 of the Fall 2016 Exam S all test time series forecasting for AR models, and Questions 21, 23, and 24 of the Fall 2017 Exam S all test the concepts of Type I and II errors.

4. Most answer choices are in the form of ranges, e.g.:

   A. Less than 1%
   B. At least 1%, but less than 2%
   C. At least 2%, but less than 3%
   D. At least 3%, but less than 4%
   E. At least 4%

   If your answer is much lower than the bound indicated by Answer A or much higher than that suggested by Answer E, do check your calculations. Chances are that you have made computational mistakes, but this is not definitely the case (sometimes the CAS examiners themselves made a mistake!).

   Note that unlike other multiple-choice exams you took before, a guessing adjustment will be in place in Exam MAS-I, so unless you can eliminate two or three of the answer choices, it will be wise of you not to answer questions which you are unsure of by pure guesswork.

What is Special about This Study Manual?

We fully understand that you have an acutely limited amount of time for study and that the exam syllabus is insanely broad. With this in mind, the overriding objective of this study manual is to help you grasp the material in Exam MAS-I, which is a new and challenging exam, effectively and efficiently, and *pass it with considerable ease*. Here are some of the invaluable features of this manual for achieving this all-important goal:

- Each chapter and section starts by explicitly stating which learning objectives and outcomes of the MAS-I exam syllabus we are going to cover, to assure you that we are on track and hitting the right target.

- The knowledge statements of the syllabus are demystified by precise and concise expositions synthesized from the syllabus readings, helping you acquire a deep and solid understanding of the subject matter.

- Formulas and results of utmost importance are $\boxed{\text{boxed}}$ for easy identification and memorization.

- To succeed in any (actuarial) exam, the importance of practicing a wide variety of non-trivial problems to sharpen your understanding and to develop proficiency, as always, cannot be overemphasized. This study manual embraces this learning by doing approach and intersperses its expositions with more than **450 in-text examples** and **720 end-of-chapter/section problems** (the harder ones are labeled as [**HARDER!**] or [**VERY HARD!!**]), which are original or taken from relevant SOA/CAS past exams, all with step-by-step solutions and problem-solving remarks, to consolidate your understanding and to give you a sense of what you can expect to see in the real exam. As you read this manual, skills are honed and confidence is built. As a general guide, you should *study all of the in-text examples with particular attention paid to recent Exam MAS-I and S questions and work out at least half of the end-of-chapter/section problems.*

- While the focus of this study manual is on exam preparation, we will not shy away from explaining the meaning of various formulas in the syllabus. The interpretations and insights provided will foster a genuine understanding of the syllabus material and discourage slavish memorization. At times, we will present brief derivations in the hope that they can help you appreciate the structure of the formulas in question. It is the author's belief and personal experience that a solid understanding of the underlying concepts is always conducive to achieving good exam results.

- Mnemonics and shortcuts are emphasized, so are highlights of important exam items and common mistakes committed by students.

- Three full-length practice exams updated for the MAS-I exam syllabus and designed to mimic the real exam conclude this study manual giving you a holistic review of the syllabus material.

New to the Fall 2018 Edition

- Old SOA/CAS exam questions before 2000, which are not easily available nowadays, are added as appropriate. Despite the seniority of these past exam questions and that different syllabus texts were used when these exams were offered, they are by no means obsolete and will prove instrumental in illustrating some less commonly tested concepts in the current syllabus and consolidating your understanding as you progress along this manual.

- All of the 45 questions from the very recent Spring 2018 Exam MAS-I are inserted into this manual and carefully discussed and solved. Variants of some of these exam questions are developed.

- A number of sections have been substantially revised, partly in response to the recent Exam S and MAS-I papers, e.g., Section 10.2 on model construction, Section 10.5 on prediction, Section 11.5 on dimension reduction methods, Section 12.2 on the method of scoring. A number of new examples and end-of-chapter/section problems have been added. *In Bayesian parlance, we learn from experience as life moves on!*

- All known typographical errors have been fixed.

Exam Tables

In the exam, you will be supplied with a variety of tables, including:

- *Standard normal distribution table* (used throughout this study manual)

  You will need this table for values of the standard normal distribution function or standard normal quantiles, when you work with normally distributed random variables or perform normal approximation.

- *Illustrative Life Table* (used mostly in Chapter 4 of this study manual)

  You will need this when you are told that mortality of the underlying population follows the Illustrative Life Table.

- *A table of distributions for a number of common continuous and discrete distributions and the formulas for their moments and other probabilistic quantities* (used throughout Parts I and II of this study manual)

  This big table provides a great deal of information about some common as well as non-common distributions (e.g., inverse exponential, inverse Gaussian, Pareto, Burr, etc.). When an exam question centers on these distributions and quantities such as their means or variances are needed, consult this table.

- *Quantiles of t-distribution, F-distribution, chi-square distribution* (used in Chapters 8, 10, 12 and 13 of this study manual)

  These quantiles will be of use when you perform parametric hypothesis tests.

You should download these tables from `http://www.casact.org/admissions/syllabus/MASI_Tables.pdf` right away, print out a copy and learn how to locate the relevant entries in these tables because they will be intensively used during your study as well as in the exam.

Acknowledgment

I would like to thank my colleagues, Professor Elias S. W. Shiu and Dr. Michelle A. Larson, at the University of Iowa for sharing with me many pre-2000 SOA/CAS exam papers. These hard-earned old exam papers have proved invaluable in illustrating a number of less commonly tested exam topics. Thanks are also due to Mr. Zhaofeng Tang, doctoral student in actuarial science at the University of Iowa, for his professional assistance in the production of some of the graphs in this study manual.

Errata

While we go to great lengths to polish and proofread this manual, some mistakes will inevitably go unnoticed. The author wishes to apologize in advance for any errors, typographical or otherwise, and would greatly appreciate it if you could bring them to his attention by *sending any errors you identify to* `ambrose-lo@uiowa.edu` *and c.c.* `support@actexmadriver.com`. Compliments and criticisms are also welcome. The author will try his best to respond to any inquiries within 48 hours and an ongoing errata list will be maintained online at `https://sites.google.com/site/ambroseloyp/publications/MAS-I`. More importantly, *students who report errors will quality for a quarterly drawing for a $100 in-store credit.*

Ambrose Lo
June 2018
Iowa City, IA

About the Author

Professor Ambrose Lo was born, raised, and educated in Hong Kong. He earned his B.S. in Actuarial Science (first class honors) and Ph.D. in Actuarial Science from The University of Hong Kong in 2010 and 2014, respectively. He joined the Department of Statistics and Actuarial Science at The University of Iowa in August 2014 as an Assistant Professor in Actuarial Science. He is a Fellow of the Society of Actuaries (FSA) and a Chartered Enterprise Risk Analyst (CERA). His research interests lie in dependence structures, quantitative risk management as well as optimal (re)insurance. His research papers have been published in top-tier actuarial journals, such as *ASTIN Bulletin: The Journal of the International Actuarial Association*, *Insurance: Mathematics and Economics*, and *Scandinavian Actuarial Journal*.

Besides dedicating himself to actuarial research, Ambrose attaches equal importance to teaching, through which he nurtures the next generation of actuaries and serves the actuarial profession. He has taught courses on financial derivatives, mathematical finance, life contingencies, credibility theory, advanced probability theory, and regression analysis. His emphasis in teaching is always placed on thorough understanding of the subject matter complemented by concrete problem-solving skills. As a result of his exceptional teaching performance, Ambrose has won numerous teaching awards ever since he was a graduate student (see `http://www.scifac.hku.hk/news/comm/award-excellence-teaching-assistant-2011-12-2` for example). He is also the coauthor of the *ACTEX Study Manual for SOA Exam SRM* (September 2018 Edition) and the sole author of the textbook *Derivative Pricing: A Problem-Based Primer*.

Part I

Probability Models (Stochastic Processes and Survival Models)

Chapter 1

Poisson Processes

LEARNING OBJECTIVES

1. Understand and apply the properties of Poisson processes:
   - For increments in the homogeneous case
   - For interval times in the homogeneous case
   - For increments in the non-homogeneous case
   - Resulting from special types of events in the Poisson process
   - Resulting from sums of independent Poisson processes

   Range of weight: 0-5 percent
2. For any Poisson process and the inter-arrival and waiting distributions associated with the Poisson process, calculate:
   - Expected values
   - Variances
   - Probabilities

   Range of weight: 0-5 percent
3. For a compound Poisson process, calculate moments associated with the value of the process at a given time.

   Range of weight: 0-5 percent
4. Apply the Poisson Process concepts to calculate the hazard function and related survival model concepts.
   - Relationship between hazard rate, probability density function and cumulative distribution function
   - Effect of memoryless nature of Poisson distribution on survival time estimation

   Range of weight: 2-8 percent

Chapter overview: As a prospective P&C actuary, you would be interested in monitoring the number of insurance claims an insurance company receives as time goes by and how these claims can be appropriately analyzed by means of sound statistical analysis. In Exam MAS-I, we shall learn one way of modeling the flow of insurance claims – the Poisson process.

This part of the syllabus has two required readings:

(1) *A study note by J.W. Daniel*

The study note is precise and concise, introducing main results mostly without proof and supplementing its exposition with a few simple examples. It is suitable for a first-time introduction to Poisson processes.

(2) *The book entitled Introduction to Probability Models by S.M. Ross.*

This is a textbook used by a number of college courses on elementary applied probability. It balances rigor and intuition, and presents the theory of Poisson processes at a level that is much deeper than that in the study note by Daniel. In particular, it treats the conditional distribution of the arrival times as well as the interplay between two independent Poisson processes. A conspicuous feature of this book is its large number of sophisticated examples and exercises which require a large amount of ingenuity and cannot be done in a reasonable exam setting. This study manual improves the practicality of the book and rephrases these otherwise intractable examples in an exam tone.

You can download this book (Eleventh Edition) "legally" from *ScienceDirect* via the following link, chapter by chapter, if your university has subscribed to it:

`http://www.sciencedirect.com/science/book/9780124079489`

You should do so because the book has a number of good exercises which will be solved in full in this study manual (the questions cannot be reproduced here because of copyright issues). The exercises had been the theme of some past SOA/CAS examination questions, so you should not despise these exercises as irrelevant and useless.

The Daniel study note has been on the syllabuses of Exams 3, 3L and ST, whereas *Introduction to Probability Models* just entered the syllabus of Exam S in Fall 2015. As a result of the addition of the latter reading, we expect more complex exam questions on Poisson processes in Exam MAS-I. In total, expect about 4 questions on the material of the entire chapter.

1.1 Fundamental Properties

KNOWLEDGE STATEMENTS

1a. Poisson process

1b. Non-homogeneous Poisson process

OPTIONAL SYLLABUS READING(S)

- Ross, Subsections 5.3.1 and 5.3.2
- Daniel, Section 1.1 and Subsection 1.4.1

Definition. By definition, a *Poisson process* $\{N(t),\, t \geq 0\}$ with *rate function* (also known as *intensity function*) $\lambda(\cdot)$ [i] is a stochastic process, namely, a collection of random variables indexed by time t (in an appropriate unit, e.g., minute, hour, month, year, etc.), satisfying the following properties:

1. *(Counting)* $N(0) = 0$, $N(t)$ is non-decreasing in t and takes non-negative integer values only.

   Interpretation: $N(t)$ counts the number of claims which are submitted on or before time t. Thus $N(0)$ is 0 (we assume that should be no claims before the insurance company starts its business), $N(t)$ cannot decrease in time and must be integer-valued.

2. *(Distribution of increments are Poisson random variables)* For $s < t$, the increment $N(t) - N(s)$, which counts the number of events in the interval $(s, t]$, is a Poisson random variable with mean $\Lambda = \int_s^t \lambda(y)\,\mathrm{d}y$.

   Interpretation: Increments of a Poisson process, as its name suggests, are Poisson random variables with mean computed by integrating the rate function over the same interval. In this regard, we can see that the rate function of a Poisson process completely specifies the distribution of each increment.

3. *(Increments are independent)* If $(s_1, t_1]$ and $(s_2, t_2]$ are non-overlapping intervals, then $N(t_1) - N(s_1)$ and $N(t_2) - N(s_2)$ are independent random variables.

   Interpretation: This is the most amazing property of a Poisson process. Its increments not only follow Poisson distribution, but also are independent on disjoint intervals (e.g., $(0, 1)$ and $(2, 5)$ are disjoint intervals, so are $(3, 4]$ and $(4, 5]$). This means that, in this model, the frequency of claims you received last month has nothing to do with the frequency this month.

[i] The study note by Daniel simply writes a Poisson process as N in short. While this is a perfectly correct way of writing, some students may confuse that with a Poisson random variable N. Also, here we write $\lambda(\cdot)$ with a parenthesis containing the argument of the function instead of just λ to emphasize that $\lambda(\cdot)$ is a function.

In the context of insurance applications, we interpret $N(t)$ as the number of claims that occur on or before time t. The same interpretation can easily carry over to more general contexts where we are interested in counting a particular type of event, e.g., the number of customers that enter a store, the number of cars passing through an intersection, the number of lucky candidates passing Exam MAS-I, etc.

Homogeneous Poisson processes. A Poisson process whose rate function is constant, say $\lambda(t) = \lambda$ for all $t \geq 0$, is called a *homogeneous* Poisson process. In addition to having independent increments, a homogeneous Poisson process also possesses *stationary increments*, meaning that the distribution of $N(t+s) - N(s)$ depends only on the length of the interval, which is t in this case, but not on s.

Probability calculations. The second and third properties of a Poisson process allow us to calculate many probabilistic quantities, such as the probability of a certain number of events, as well as the expected and variance of the number of events in a particular time interval. These two properties will be intensively used in exam questions. The following string of past exam questions serves as excellent illustrations.

RECALL

Just in case you forgot:

1. The probability mass function of a Poisson random variable X with parameter λ (a scalar, not a function) is given by

$$\Pr(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

   The mean and variance of X are both equal to λ.

2. If $X_1, X_2, \ldots, X_n$ are independent Poisson random variables with respective means $\lambda_1, \lambda_2, \ldots, \lambda_n$, then $X_1+X_2+\cdots+X_n$ is also a Poisson random variable with a mean of $\lambda_1 + \lambda_2 + \cdots + \lambda_n$. In other words, the sum of independent Poisson random variables is also a Poisson random variable whose mean is the sum of the individual Poisson means.

Example 1.1.1. (SOA/CAS Exam P/1 Sample Question 173: Warm-up question) In a given region, the number of tornadoes in a one-week period is modeled by a Poisson distribution with mean 2. The numbers of tornadoes in different weeks are mutually independent.

Calculate the probability that fewer than four tornadoes occur in a three-week period.

A. 0.13

B. 0.15

C. 0.29

D. 0.43

E. 0.86

Ambrose's comments: This is not even a Poisson process question. It simply reminds you of how probabilities for a Poisson random variable are typically calculated.

Solution. We are interested in $\Pr(N_1+N_2+N_3 < 4)$, where N_i is the number of tornadoes in the i^{th} week for $i = 1, 2, 3$. As $N_1 + N_2 + N_3$ is also a Poisson random variable with a mean of $3(2) = 6$, we have

$$\begin{aligned}
\Pr(N_1 + N_2 + N_3 < 4) &= \sum_{i=0}^{3} \Pr(N_1 + N_2 + N_3 = i) \\
&= \mathrm{e}^{-6}\left(1 + 6 + \frac{6^2}{2!} + \frac{6^3}{3!}\right) \\
&= \boxed{0.1512}. \quad \textbf{(Answer: B)}
\end{aligned}$$

□

Example 1.1.2. (CAS Exam MAS-I Spring 2018 Question 3: Probability – I) The number of cars passing through the Lexington Tunnel follows a Poisson process with rate:

$$\lambda(t) = \begin{cases} 16 + 2.5t & \text{for } 0 < t \leq 8 \\ 52 - 2t & \text{for } 8 < t \leq 12 \\ -20 + 4t & \text{for } 12 < t \leq 18 \\ 160 - 6t & \text{for } 18 < t \leq 24 \end{cases}$$

Calculate the probability that exactly 50 cars pass through the tunnel between times $t = 11$ and $t = 13$.

A. Less than 0.01

B. At least 0.01, but less than 0.02

C. At least 0.02, but less than 0.03

D. At least 0.03, but less than 0.04

E. At least 0.04

Solution. The number of cars between $t = 11$ and $t = 13$ is a Poisson random variable with parameter

$$\begin{aligned}\int_{11}^{13} \lambda(t)\,\mathrm{d}t &= \int_{11}^{12} (52 - 2t)\,\mathrm{d}t + \int_{12}^{13} (-20 + 4t)\,\mathrm{d}t \\ &= [52 - (12^2 - 11^2)] + [-20 + 2(13^2 - 12^2)] \\ &= 59.\end{aligned}$$

Hence the probability of exactly 50 cars between $t = 11$ and $t = 13$ is

$$\frac{\mathrm{e}^{-59}59^{50}}{50!} = \boxed{0.0273}. \qquad \textbf{(Answer: C)}$$

□

Example 1.1.3. (CAS Exam 3L Spring 2010 Question 12: Probability – II) Downloads of a song on a musician's Web site follow a heterogeneous Poisson process with the following Poisson rate function:

$$\lambda(t) = \mathrm{e}^{-0.25t}$$

Calculate the probability that there will be more than two downloads of this song between times $t = 1$ and $t = 5$.

A. Less than 29%

B. At least 29%, but less than 30%

C. At least 30%, but less than 31%

D. At least 31%, but less than 32%

E. At least 32%

Solution. Because $N(5) - N(1)$ is a Poisson random variable with parameter

$$\int_1^5 \lambda(t)\,\mathrm{d}t = \int_1^5 \mathrm{e}^{-0.25t}\,\mathrm{d}t = \frac{\mathrm{e}^{-0.25(1)} - \mathrm{e}^{-0.25(5)}}{0.25} = 1.969184,$$

the required probability equals

$$\begin{aligned}&\Pr(N(5) - N(1) > 2) \\ =\ & 1 - \Pr(N(5) - N(1) = 0) - \Pr(N(5) - N(1) = 1) - \Pr(N(5) - N(1) = 2) \\ =\ & 1 - \mathrm{e}^{-1.969184}\left(1 + 1.969184 + \frac{1.969184^2}{2}\right) \\ =\ & \boxed{0.3150}. \quad \textbf{(Answer: D)}\end{aligned}$$

□

Example 1.1.4. (CAS Exam S Spring 2016 Question 3: Probability – III) You are given:

- The number of claims, $N(t)$, follows a Poisson process with intensity:

$$\begin{aligned}\lambda(t) &= \frac{1}{2}t, \quad 0 < t < 5\\ \lambda(t) &= \frac{1}{4}t, \quad t \geq 5\end{aligned}$$

- By time $t = 4$, 15 claims have occurred.

Calculate the probability that exactly 16 claims will have occurred by time $t = 6$.

A. Less than 0.075

B. At least 0.075, but less than 0.125

C. At least 0.125, but less than 0.175

D. At least 0.175, but less than 0.225

E. At least 0.225

Solution. The number of claims between $t = 4$ and $t = 6$ is a Poisson random variable with mean

$$\int_4^5 \frac{1}{2}t\,\mathrm{d}t + \int_5^6 \frac{1}{4}t\,\mathrm{d}t = \frac{5^2-4^2}{2(2)} + \frac{6^2-5^2}{4(2)} = 3.625.$$

The probability of having exactly one($= 16 - 15$) claim between $t = 4$ and $t = 6$ is

$$3.625\mathrm{e}^{-3.625} = \boxed{0.0966}. \quad \textbf{(Answer: B)}$$

□

Remark. More formally, the probability we seek is

$$\begin{aligned}\Pr(N(6) = 16|N(4) = 15) &= \Pr(N(6) - N(4) = 1|N(4) = 15)\\ &= \Pr(N(6) - N(4) = 1)\end{aligned}$$

due to the property of independent increments.

Example 1.1.5. (CAS Exam 3L Spring 2012 Question 9: Expected value from now until forever) Claims reported for a group of policies follow a non-homogeneous

Poisson process with rate function:

$$\lambda(t) = 100/(1+t)^3, \text{ where } t \text{ is the time (in years) after January 1, 2011.}$$

Calculate the expected number of claims reported after January 1, 2011 for this group of policies.

A. Less than 45

B. At least 45, but less than 55

C. At least 55, but less than 65

D. At least 65, but less than 75

E. At least 75

Solution. We are interested in $N(\infty) = \lim_{t\to\infty} N(t)$, which is a Poisson random variable with mean

$$\int_0^\infty \lambda(t)\,\mathrm{d}t = \int_0^\infty \frac{100}{(1+t)^3}\,\mathrm{d}t = 100\left[-\frac{1}{2(1+t)^2}\right]_0^\infty = \boxed{50}. \quad \textbf{(Answer: B)}$$

□

Example 1.1.6. (CAS Exam 3L Spring 2013 Question 9: Variance) You are given the following:

- An actuary takes a vacation where he will not have access to email for eight days.
- While he is away, emails arrive in the actuary's inbox following a non-homogeneous Poisson process where

$$\lambda(t) = 8t - t^2 \text{ for } 0 \le t \le 8. \, (t \text{ is in days})$$

Calculate the variance of the number of emails received by the actuary during this trip.

A. Less than 60

B. At least 60, but less than 70

C. At least 70, but less than 80

D. At least 80, but less than 90

E. At least 90

Solution. The trip of the actuary lasts for 8 days, during which the number of emails is a Poisson random variable with variance (same as the Poisson parameter)

$$\int_0^8 (8t - t^2)\,\mathrm{d}t = \left[4t^2 - \frac{t^3}{3}\right]_0^8 = \boxed{85.3333}. \quad \textbf{(Answer: D)}$$

□

Example 1.1.7. (CAS Exam ST Fall 2015 Question 1: Calculation of homogeneous Poisson intensity) For two Poisson processes, N_1 and N_2, you are given:

- N_1 has intensity function $\lambda_1(t) = \begin{cases} 2t & \text{for } 0 < t \le 1 \\ t^3 & \text{for } t > 1 \end{cases}$
- N_2 is a homogeneous Poisson process.
- $\mathrm{Var}[N_1(3)] = 4\mathrm{Var}[N_2(3)]$

Calculate the intensity of N_2 at $t = 3$.

A. Less than 1

B. At least 1, but less than 3

C. At least 3, but less than 5

D. At least 5, but less than 7

E. At least 7

Solution. Note that $N_1(3)$ has a mean and variance equal to

$$\int_0^3 \lambda_1(t)\,\mathrm{d}t = \int_0^1 2t\,\mathrm{d}t + \int_1^3 t^3\,\mathrm{d}t = [t^2]_0^1 + \left[\frac{t^4}{4}\right]_1^3 = 1 + \frac{3^4 - 1^4}{4} = 21,$$

while $N_2(3)$ has a mean and variance equal to $3\lambda_2$, where λ_2 is the constant intensity of N_2. As $\mathrm{Var}[N_1(3)] = 4\mathrm{Var}[N_2(3)]$, we have $21 = 4(3\lambda_2)$, so $\lambda_2 = 21/12 = \boxed{1.75}$. **(Answer: B)** □

Probabilities involving overlapping intervals. A harder exam question may ask that you determine probabilities for increments on overlapping intervals. The key step to calculate these probabilities lies in rewriting the events in terms of increments on *non-overlapping* intervals, which are independent according to the definition of a Poisson process.

Example 1.1.8. (Probability for overlapping increments I) The number of calls received in a telephone exchange follow a homogeneous Poisson process with a rate of 30 per hour.

Calculate the probability that there are exactly 2 calls in the first ten minutes and exactly 5 calls in the first twenty minutes.

A. Less than 0.01

B. At least 0.01, but less than 0.02

C. At least 0.02, but less than 0.03

D. At least 0.03, but less than 0.04

E. At least 0.04

Solution. When time is measured in hours, the required probability is

$$\Pr(N(1/6) = 2, N(1/3) = 5) = \Pr(N(1/6) = 2, N(1/3) - N(1/6) = 3),$$

which can be factored, because of independence, into

$$\begin{aligned} \Pr(N(1/6) = 2) \Pr(N(1/3) - N(1/6) = 3) &= \frac{\mathrm{e}^{-30/6}(30/6)^2}{2!} \times \frac{\mathrm{e}^{-30/6}(30/6)^3}{3!} \\ &= \boxed{0.0118}. \quad \textbf{(Answer: B)} \end{aligned}$$

□

Example 1.1.9. (Probability for overlapping increments II) Customers arrive at a post office in accordance with a Poisson process with a rate of 5 per hour. The post office opens at 9:00 am.

Calculate the probability that only one customer arrives before 9:20 am and ten customers arrive before 11:20 am.

A. Less than 0.01

B. At least 0.01, but less than 0.02

C. At least 0.02, but less than 0.03

D. At least 0.03, but less than 0.04

E. At least 0.04

Solution. The required probability is

$$\begin{aligned}
\Pr\left(N(1/3)=1, N(7/3)=10\right) &= \Pr\left(N(1/3)=1, N(7/3)-N(1/3)=9\right)\\
&= \Pr\left(N(1/3)=1\right)\Pr\left(N(7/3)-N(1/3)=9\right)\\
&= \mathrm{e}^{-5/3}\left(\frac{5}{3}\right)\times \mathrm{e}^{-5(2)}\frac{[5(2)]^9}{9!}\\
&= \boxed{0.0394}. \quad \textbf{(Answer: D)}
\end{aligned}$$

□

Conditional distribution of $N(t)$ given $N(s)$ for $s \le t$. Suppose that we know the value of the Poisson process at one time point s with $N(s) = m$, and we wish to study the probabilistic behavior of the Poisson process at a later time point t with $s \le t$. Then $N(t)$ turns out to be a *translated Poisson* random variable in the sense that it has the same distribution as the sum of a Poisson random variable and a constant. To see this, let's write $N(t)$ as

$$N(t) = [N(t) - N(s)] + N(s).$$

The second term $N(s)$ is known to be m, while the first term, owing to the property of independent increments of a Poisson process, is a Poisson random variable, say M, whose distribution does not depend on the value of m. Therefore, we have the distributional representation

$$\boxed{[N(t)|N(s)=m] \stackrel{d}{=} M + m, \quad s \le t,}$$

where "$\stackrel{d}{=}$" means equality in distribution. This result allows us to answer questions about many probabilistic quantities associated with $N(t)$ when the value of $N(s)$ is given.

Example 1.1.10. (CAS Exam S Fall 2017 Question 3: Conditional probability) You are given:

- A Poisson process N has a rate function: $\lambda(t) = 3t^2$
- You've already observed 50 events by time $t = 2.1$.

Calculate the conditional probability, $\Pr[N(3) = 68 \mid N(2.1) = 50]$.

A. Less than 5%

B. At least 5%, but less than 10%

C. At least 10%, but less than 15%

D. At least 15%, but less than 20%

E. At least 20%

Solution. The conditional probability can be determined as

$$\begin{aligned}\Pr[N(3) = 68 \mid N(2.1) = 50] &= \Pr[N(3) - N(2.1) = 18 \mid N(2.1) = 50] \\ &= \Pr[N(3) - N(2.1) = 18],\end{aligned}$$

where $N(3) - N(2.1)$ is a Poisson random variable with mean $\int_{2.1}^{3} 3t^2 \,\mathrm{d}t = 3^3 - 2.1^3 = 17.739$. The final answer is

$$\frac{\mathrm{e}^{-17.739} 17.739^{18}}{18!} = \boxed{0.0934}. \quad \textbf{(Answer: B)}$$

□

Example 1.1.11. (CAS Exam 3L Fall 2010 Question 11: Conditional variance) You are given the following information:

- A Poisson process N has a rate function $\lambda(t) = 3t^2$.
- You have observed 50 events by time $t = 2.1$.

Calculate $\mathrm{Var}[N(3) \mid N(2.1) = 50]$.

A. Less than 10

B. At least 10, but less than 20

C. At least 20, but less than 30

D. At least 30, but less than 40

E. At least 40

Solution. Conditional on $N(2.1) = 50$, $N(3)$ has the same distribution as $M + 50$, where M is a Poisson random variable with mean and variance

$$\int_{2.1}^{3} \lambda(t) \,\mathrm{d}t = \int_{2.1}^{3} 3t^2 \,\mathrm{d}t = t^3\Big|_{2.1}^{3} = 17.739.$$

Hence

$$\mathrm{Var}[N(3) \mid N(2.1) = 50] = \mathrm{Var}(M + 50) = \mathrm{Var}(M) = \boxed{17.739} \quad \textbf{(Answer: B)}.$$

□

Normal approximation. For more cumbersome probabilities such as $\Pr(N(t) > c)$ with c being a large number, exact calculations can be tedious and normal approximation may be used. That is, we approximate $N(t)$ by a normal random variable with the same mean and variance, and instead calculate the probability of the same event for this normal random variable using the standard normal distribution table you have in the exam. Because the distribution of $N(t)$ is discrete, and a continuous distribution (i.e., normal) is used to approximate this discrete distribution, a continuity correction should be made.

Recall - Continuity correction

Let X be a random variable taking values in the set of integers $\{0, \pm 1, \pm 2, \ldots\}$ and N is a normal random variable having the same mean and variance as X. The following shows how various probabilities are approximated using the normal approximation *with* continuity correction: (c is an integer)

Probability of Interest		Approximant
$\Pr(X \le c)$	$\approx$	$\Pr(N \le c + 0.5)$
$\Pr(X < c)$	$\approx$	$\Pr(N \le c - 0.5)$
$\Pr(X \ge c)$	$\approx$	$\Pr(N \ge c - 0.5)$
$\Pr(X > c)$	$\approx$	$\Pr(N > c + 0.5)$

In the second column, it does not matter whether we take strict or weak inequalities because N is a continuous random variable. In other words, we may replace "$\le$" by "$<$" and "$\ge$" by "$>$".

Throughout this study manual, we denote the distribution function of the standard normal distribution by Φ.

Example 1.1.12. (CAS Exam 3L Spring 2008 Question 11: Normal approximation) A customer service call center operates from 9:00 AM to 5:00 PM. The number of calls received by the call center follows a Poisson process whose rate function varies according to the time of day, as follows:

Time of Day	Call Rate (per hour)
9:00 AM to 12:00 PM	30
12:00 PM to 1 :00 PM	10
1:00 PM to 3:00 PM	25
3:00 PM to 5:00 PM	30

Using a normal approximation, what is the probability that the number of calls received from 9:00AM to 1:00PM exceeds the number of calls received from 1:00PM to 5:00PM?

A. Less than 10%

B. At least 10%, but less than 20%

C. At least 20%, but less than 30%

D. At least 30%, but less than 40%

E. At least 40%

Solution. The number of calls received from 9:00AM to 1:00PM is a Poisson random variable N^1 with parameter $30(3)+10(1)=100$, while the number of calls received from 1:00PM to 5:00PM is a Poisson random variable N^2 with parameter $25(2)+30(2)=110$. Because N^1 and N^2 are independent,

$$\mathrm{E}[N^1-N^2]=\mathrm{E}[N^1]-\mathrm{E}[N^2]=100-110=-10,$$

and

$$\mathrm{Var}(N^1-N^2)=\mathrm{Var}(N^1)+\mathrm{Var}(N^2)=100+110=210.$$

Using the normal approximation with continuity correction, we have

$$\Pr(N^1>N^2)=\Pr(N^1-N^2>0)\approx\Pr\Bigg(\underbrace{\mathrm{N}(-10,210)}_{\substack{\text{a normal r.v. with mean}\\ -10\text{ and variance }210}}>0.5\Bigg),$$

which, upon standardization, equals

$$\Pr\left(\mathrm{N}(0,1)>\frac{0.5-(-10)}{\sqrt{210}}\right)=1-\Phi(0.72)=1-0.7642=\boxed{0.2358}. \quad \textbf{(Answer: C)}$$

□

Remark. If you do not use continuity correction, you will calculate

$$\begin{aligned}\Pr(N^1>N^2) &\approx \Pr\left(\mathrm{N}(-10,210)>0\right)\\ &= 1-\Phi\left(\frac{0-(-10)}{\sqrt{210}}\right)=1-\underbrace{\Phi(0.69)}_{0.7549}=0.2451,\end{aligned}$$

in which case you will also end up with Answer C.

[HARDER!] Conditional distribution of $N(s)$ given $N(t)$ with $s\le t$. We have learned that conditional on $N(s)$, the distribution of $N(t)$, where $0\le s\le t$, is that of a translated Poisson distribution. What about the conditional distribution of $N(s)$ given

$N(t)$? To answer this question, we consider, for $k = 0, 1, \ldots, n$,

$$\begin{aligned}\Pr(N(s) = k | N(t) = n) &= \frac{\Pr(N(s) = k, N(t) = n)}{\Pr(N(t) = n)} \\ &= \frac{\Pr(N(s) = k, N(t) - N(s) = n - k)}{\Pr(N(t) = n)}.\end{aligned}$$

Because a (homogeneous or non-homogeneous) Poisson process possesses independent increments, the preceding probability can be further written as

$$\begin{aligned}\Pr(N(s) = k | N(t) = n) &= \frac{\Pr(N(s) = k)\Pr(N(t) - N(s) = n - k)}{\Pr(N(t) = n)} \\ &= \frac{\mathrm{e}^{-m(s)}[m(s)]^k/k! \times \mathrm{e}^{-[m(t)-m(s)]}[m(t) - m(s)]^{n-k}/(n-k)!}{\mathrm{e}^{-m(t)}[m(t)]^n/n!} \\ &= \frac{n!}{k!(n-k)!}\left(\frac{m(s)}{m(t)}\right)^k \left(1 - \frac{m(s)}{m(t)}\right)^{n-k} \\ &= \binom{n}{k}\left(\frac{m(s)}{m(t)}\right)^k \left(1 - \frac{m(s)}{m(t)}\right)^{n-k}.\end{aligned}$$

In other words, given $N(t) = n$, $N(s)$ is a binomial random variable with parameters n and $m(s)/m(t)$. In particular, for a homogeneous Poisson process with rate λ, i.e., $m(t) = \lambda t$ for $t \geq 0$, then

$$\boxed{N(s)|N(t) = n \quad \sim \quad \text{Binomial}\left(n, \frac{s}{t}\right),}$$

which is free of λ.

Example 1.1.13. (CAS Exam MAS-I Spring 2018 Question 2: Probability for $N(s)$ given $N(t)$ with $s \leq t$) Insurance claims are made according to a Poisson process with rate λ.

Calculate the probability that exactly 3 claims were made by time $t = 1$, given that exactly 6 claims are made by time $t = 2$.

A. Less than 0.3

B. At least 0.3, but less than 0.4

C. At least 0.4, but less than 0.5

D. At least 0.5, but less than 0.6

E. At least 0.6

Solution. Conditional on $N(2) = 6$, $N(1)$ is a binomial random variable with parameters 6 and 1/2, so

$$\Pr[N(1) = 3|N(2) = 6] = \binom{6}{3}\left(\frac{1}{2}\right)^3\left(1-\frac{1}{2}\right)^3 = \boxed{0.3125}. \quad \textbf{(Answer: B)}$$

□

Remark. We do not need the value of λ to get the answer.

Problems

Problem 1.1.1. (SOA Exam P Sample Question 280: Conditional Poisson mean) The number of burglaries occurring on Burlington Street during a one-year period is Poisson distributed with mean 1.

Calculate the expected number of burglaries on Burlington Street in a one-year period, given that there are at least two burglaries.

A. 0.63

B. 2.39

C. 2.54

D. 3.00

E. 3.78

Solution. Let N be the number of burglaries on Burlington Street in a specified one-year period. Given that there are at least two burglaries, the expected value of N is

$$\begin{aligned} \mathrm{E}[N|N \geq 2] &= \frac{\sum_{n=2}^{\infty} n \Pr(N = n)}{\Pr(N \geq 2)} \\ &= \frac{\mathrm{E}[N] - 1 \times \Pr(N = 1)}{1 - \Pr(N = 0) - \Pr(N = 1)} \\ &= \frac{1 - \mathrm{e}^{-1}}{1 - \mathrm{e}^{-1} - \mathrm{e}^{-1}} \\ &= \boxed{2.3922}. \quad \textbf{(Answer: B)} \end{aligned}$$

□

Problem 1.1.2. (CAS Exam 3 Fall 2006 Question 26: True-of-false questions) Which of the following is/are true?

1. A counting process is said to possess independent increments if the number of events that occur between time s and t is independent of the number of events that occur between time s and $t+u$ for all $u > 0$.
2. All Poisson processes have stationary and independent increments.
3. The assumption of stationary and independent increments is essentially equivalent to asserting that at any point in time the process probabilistically restarts itself.

A. 1 only

B. 2 only

C. 3 only

D. 1 and 2 only

E. 2 and 3 only

Solution. Only 3. is correct. **(Answer: C)**

1. This would be true if the second s is changed to t.
2. A non-homogeneous Poisson process does not have stationary increments in general.

□

Problem 1.1.3. (CAS Exam 3L Fall 2013 Question 9: Polynomial intensity function) You are given that claim counts follow a non-homogeneous Poisson Process with $\lambda(t) = 30t^2 + t^3$.

Calculate the probability of at least two claims between time 0.2 and 0.3.

A. Less than 1%

B. At least 1%, but less than 2%

C. At least 2%, but less than 3%

D. At least 3%, but less than 4%

E. At least 4%

Solution. The number of claims between times 0.2 and 0.3 is a Poisson random variable with parameter

$$\int_{0.2}^{0.3} (30t^2 + t^3)\,\mathrm{d}t = 10t^3 + \frac{t^4}{4}\bigg|_{0.2}^{0.3} = 0.191625.$$

Hence the probability of at least two claims between times 0.2 and 0.3 is the complement of the probability of having 0 or 1 claim:

$$1 - \Pr(0 \text{ claim}) - \Pr(1 \text{ claim}) = 1 - \mathrm{e}^{-0.191625}(1 + 0.191625) = \boxed{0.0162}. \quad \textbf{(Answer: B)}$$

□

Problem 1.1.4. (CAS Exam 3 Fall 2006 Question 28: Piecewise linear intensity function) Customers arrive to buy lemonade according to a Poisson distribution with $\lambda(t)$, where t is time in hours, as follows:

$$\lambda(t) = \begin{cases} 2+6t & 0 \le t \le 3 \\ 20 & 3 < t \le 4 \\ 36-4t & 4 < t \le 8 \end{cases}$$

At 9:00 a.m., t is 0.

Calculate the number of customers expected to arrive between 10:00 a.m. and 2:00 p.m.

A. Less than 63

B. At least 63, but less than 65

C. At least 65, but less than 67

D. At least 67, but less than 69

E. At least 69

Solution. The expected number of customers arriving between 10:00 a.m. $(t = 1)$ and 2:00 p.m. $(t = 5)$ is

$$\begin{aligned} \int_1^5 \lambda(t)\,\mathrm{d}t &= \int_1^3 (2+6t)\,\mathrm{d}t + \int_3^4 20\,\mathrm{d}t + \int_4^5 (36-4t)\,\mathrm{d}t \\ &= [2t+3t^2]_1^3 + 20 + [36t-2t^2]_4^5 \\ &= \boxed{66}. \quad \textbf{(Answer: C)} \end{aligned}$$

□

Remark. Because the intensity function is piecewise linear, integrating it is the same as calculating the areas of trapeziums.

Problem 1.1.5. (SOA Course 3 Fall 2004 Question 26: Linear rate function) Customers arrive at a store at a Poisson rate that increases linearly from 6 per hour at 1:00 p.m. to 9 per hour at 2:00 p.m.

Calculate the probability that exactly 2 customers arrive between 1:00 p.m. and 2:00 p.m.

A. 0.016

B. 0.018

C. 0.020

D. 0.022

E. 0.024

Solution. Let 1:00 p.m. be time 0 and measure time in hours. The rate function is given by

$$\lambda(t) = 6 + 3t, \quad t \geq 0.$$

You can check that $\lambda(0) = 6$ and $\lambda(1) = 9$. The number of customers that arrive between 1:00 p.m. and 2:00 p.m. is a Poisson random variable with a mean of $\int_0^1 \lambda(t)\,\mathrm{d}t = 6 + 3/2 = 7.5$. The probability of having 2 customers in the same period is

$$\frac{\mathrm{e}^{-7.5}(7.5)^2}{2!} = \boxed{0.0156}. \quad \textbf{(Answer: A)}$$

□

Problem 1.1.6. (CAS Exam 3L Fall 2008 Question 1: Expected value) The number of accidents on a highway from 3:00 PM to 7:00 PM follows a nonhomogeneous Poisson process with rate function

$$\lambda = 4 - (t-2)^2, \text{ where } t \text{ is the number of hours since 3:00 PM.}$$

How many more accidents are expected from 4:00 PM to 5:00 PM than from 3:00 PM to 4:00PM?

A. Less than 0.75

B. At least 0.75, but less than 1.25

C. At least 1.25, but less than 1.75

D. At least 1.75, but less than 2.25

E. At least 2.25

Solution.

- The expected number of accidents from 3:00 PM to 4:00PM is

$$\int_0^1 [4 - (t-2)^2]\,\mathrm{d}t = \left[4t - \frac{(t-2)^3}{3}\right]_0^1 = \frac{5}{3}.$$

- The expected number of accidents from 4:00 PM to 5:00 PM is

$$\int_1^2 [4 - (t-2)^2]\,\mathrm{d}t = \left[4t - \frac{(t-2)^3}{3}\right]_1^2 = \frac{11}{3}.$$

The difference is $\boxed{2}$. **(Answer: D)** □

Problem 1.1.7. (CAS Exam 3L Fall 2008 Question 2: Probability, homogeneous) You are given the following:

- Hurricanes occur at a Poisson rate of 1/4 per week during the hurricane season.
- The hurricane season lasts for exactly 15 weeks.

Prior to the next hurricane season, a weather forecaster makes the statement, "There will be at least three and no more than five hurricanes in the upcoming hurricane season."

Calculate the probability that this statement will be correct.

A. Less than 54%

B. At least 54%, but less than 56%

C. At least 56%, but less than 58%

D. At least 58%, but less than 60%

E. At least 60%

Solution. Note that $N(15)$, the number of hurricanes during the 15-week hurricane season, is a Poisson random variable with a mean of $15/4 = 3.75$. The probability that the statement will be correct is

$$\Pr(3 \leq N(15) \leq 5) = \mathrm{e}^{-3.75}\left(\frac{3.75^3}{3!} + \frac{3.75^4}{4!} + \frac{3.75^5}{5!}\right) = \boxed{0.5458}. \quad \textbf{(Answer: B)}$$

□

Problem 1.1.8. (CAS Exam 3L Spring 2008 Question 10: Probability, non-homogeneous) Car accidents follow a Poisson process, as described below:

- On Monday and Friday, the expected number of accidents per day is 3.
- On Tuesday, Wednesday, and Thursday, the expected number of accidents per day is 4.
- On Saturday and Sunday, the expected number of accidents per day is 1.

Calculate the probability that exactly 18 accidents occur in a week.

A. Less than .06

B. At least .06 but less than .07

C. At least .07 but less than .08

D. At least .08 but less than .09

E. At least .09

Solution. The total number of accidents in a week is a Poisson random variable with a mean of $3(2) + 4(3) + 1(2) = 20$, so the probability of having exactly 18 accidents in a week is

$$\frac{e^{-20}20^{18}}{18!} = \boxed{0.0844}. \quad \textbf{(Answer: D)}$$

□

Problem 1.1.9. (CAS Exam 3 Spring 2006 Question 33: Probability, non-homogeneous) While on vacation, an actuarial student sets out to photograph a Jackalope and a Snipe, two animals common to the local area. A tourist information booth informs the student that daily sightings of Jackalopes and Snipes follow independent Poisson processes with intensity parameters:

$$\begin{aligned} \lambda_J(t) &= \frac{t^{1/3}}{5} \qquad \text{for Jackalopes} \\ \lambda_S(t) &= \frac{t^{1/2}}{10} \qquad \text{for Snipes} \end{aligned}$$

where: $0 \le t \le 24$ and t is the number of hours past midnight

If the student takes photographs between 1 pm and 5 pm, calculate the probability that he will take at least 1 photograph of each animal.

A. Less than 0.45

B. At least 0.45, but less than 0.60

C. At least 0.60, but less than 0.75

D. At least 0.75, but less than 0.90

E. At least 0.90

Solution. The number of Jackalopes and Snipes between 1 pm and 5 pm are Poisson random variables with respective means

$$\frac{1}{5}\int_{13}^{17} t^{1/3}\,\mathrm{d}t = \frac{3}{4(5)}(17^{4/3} - 13^{4/3}) = 1.971665$$

and

$$\frac{1}{10}\int_{13}^{17} t^{1/2}\,\mathrm{d}t = \frac{2}{3(10)}(17^{3/2} - 13^{3/2}) = 1.548042.$$

Because the two Poisson processes are independent (note: here we are not using the property of independent increments),

$$\begin{aligned} & \Pr(N_J(17) - N_J(13) \ge 1, N_S(17) - N_S(13) \ge 1) \\ = \; & \Pr(N_J(17) - N_J(13) \ge 1)\Pr(N_S(17) - N_S(13) \ge 1) \\ = \; & [1 - \Pr(N_J(17) - N_J(13) = 0)]\,[1 - \Pr(N_S(17) - N_S(13) = 0)] \\ = \; & (1 - e^{-1.971665})(1 - e^{-1.548042}) \\ = \; & \boxed{0.6777}. \quad \textbf{(Answer: C)} \end{aligned}$$

□

Problem 1.1.10. (CAS Exam 3 Fall 2005 Question 26: Probability, non-homogeneous) The number of reindeer injuries on December 24 follows a Poisson process with intensity function:

$$\lambda(t) = (t/12)^{1/2} \qquad 0 \le t \le 24, \text{ where } t \text{ is measured in hours}$$

Calculate the probability that no reindeer will be injured during the last hour of the day.

A. Less than 30%

B. At least 30%, but less than 40%

C. At least 40%, but less than 50%

D. At least 50%, but less than 60%

E. At least 60%

Solution. We need

$$\begin{aligned}\Pr(N(24) - N(23) = 0) &= \exp\left[-\int_{23}^{24} (t/12)^{1/2}\,\mathrm{d}t\right] \\ &= \exp\left[-\frac{2}{3(12)^{1/2}}(24^{3/2} - 23^{3/2})\right] \\ &= \boxed{0.24675}. \quad \textbf{(Answer: A)}\end{aligned}$$

□

Problem 1.1.11. [HARDER!] (Rate function mimics the normal density function) You are given that claim counts follow a non-homogeneous Poisson process with intensity function $\lambda(t) = \mathrm{e}^{-t^2/4}$.

Calculate the probability of at least two claims between time 1 and time 2.

A. Less than 0.10

B. At least 0.10, but less than 0.15

C. At least 0.15, but less than 0.20

D. At least 0.20, but less than 0.25

E. At least 0.25

Solution. The number of claims between time 1 and time 2 is a Poisson random variable with mean

$$\int_1^2 \lambda(t)\,\mathrm{d}t = \int_1^2 \mathrm{e}^{-t^2/4}\,\mathrm{d}t.$$

Note that the integrand resembles the density function of a normal distribution with mean 0 and variance 2, except that the normalizing constant $1/\sqrt{2\pi(2)}$ is missing. Hence

$$\begin{aligned}
\int_1^2 \mathrm{e}^{-t^2/4}\,\mathrm{d}t &= \sqrt{2\pi(2)}\int_1^2 \frac{1}{\sqrt{2\pi(2)}}\mathrm{e}^{-t^2/4}\,\mathrm{d}t \\
&= 2\sqrt{\pi}\,\Pr\left(1 < \mathrm{N}(0,2) < 2\right) \\
&= 2\sqrt{\pi}\left[\Phi\left(\frac{2}{\sqrt{2}}\right) - \Phi\left(\frac{1}{\sqrt{2}}\right)\right] \\
&= 2\sqrt{\pi}[\Phi(1.41) - \Phi(0.71)] \\
&= 2\sqrt{\pi}(0.9207 - 0.7611) \\
&= 0.565767.
\end{aligned}$$

Finally, the required probability is

$$\begin{aligned}
\Pr\left(N(2) - N(1) \geq 2\right) &= 1 - \Pr\left(N(2) - N(1) \leq 1\right) \\
&= 1 - \mathrm{e}^{-0.565767}(1 + 0.565767) \\
&= \boxed{0.1108}. \quad \textbf{(Answer: B)}
\end{aligned}$$

□

Problem 1.1.12. (CAS Exam ST Spring 2016 Question 1: Probability for $N(s)$ given $N(t)$ with $s \leq t$ – I) You are given that $N(t)$ follows the Poisson process with rate $\lambda = 2$.

Calculate $\Pr[N(2) = 3|N(5) = 7]$.

A. Less than 0.25

B. At least 0.25, but less than 0.35

C. At least 0.35, but less than 0.45

D. At least 0.45, but less than 0.55

E. At least 0.55

Solution. Conditional on $N(5) = 7$, $N(2)$ is a binomial random variable with parameters 7 and 2/5, so

$$\Pr[N(2) = 3|N(5) = 7] = \binom{7}{3}\left(\frac{2}{5}\right)^3\left(1 - \frac{2}{5}\right)^4 = \boxed{0.290304}. \quad \textbf{(Answer: B)}$$

□

Remark. We do not need the value of λ to get the answer.

Problem 1.1.13. (Probability for $N(s)$ given $N(t)$ for $s \leq t$ – II) Customers arrive at a post office in accordance with a Poisson process with a rate of 5 per hour. The post office opens at 9:00 am.

Ten customers have arrived before 11:00 am.

Calculate the probability that only two customers have arrived before 9:30 am.

A. Less than 0.15

B. At least 0.15, but less than 0.20

C. At least 0.20, but less than 0.25

D. At least 0.25, but less than 0.30

E. At least 0.30

Solution. Note that $N(0.5)|N(2) = 10$ has a binomial distribution with parameters 10 and $0.5/2 = 0.25$. The conditional probability that $N(0.5) = 2$ equals

$$\binom{10}{2} 0.25^2(1-0.25)^8 = \boxed{0.2816}. \quad \textbf{(Answer: D)}$$

□

Problem 1.1.14. [HARDER!] (Mean of a conditional sandwiched Poisson process value) You are given that $\{N(t)\}$ is a Poisson process with rate $\lambda = 2$.

Calculate the expected value of $N(3)$, conditional on $N(2) = 3$ and $N(5) = 10$.

A. Less than 5

B. At least 5, but less than 6

C. At least 6, but less than 7

D. At least 7, but less than 8

E. At least 8

Solution. We are interested in the distribution of the value of a Poisson process at a particular time point, given the process values at an *earlier* time as well as a *later* time. To reduce this two-condition setting to the one-condition setting involving $N(s)$ given $N(t)$ for $s \leq t$, we consider the *translated Poisson process* $\{N^2(t)\}_{t \geq 0}$ defined by $N^2(t) := N(2+t) - N(2)$. This translated process looks at the original Poisson process $\{N(t)\}$ from time 2 (hence the superscript "2") onward, but with values translated downward by $N(2)$ units. It is easy to conceive (and can be rigorously shown) that $\{N^2(t)\}_{t \geq 0}$ is indeed a Poisson process (see Exercise 5.35 of Ross). Moreover, because of independent increments, $\{N^2(t)\}_{t \geq 0}$ is independent of $N(2)$. In terms of the translated Poisson process, $N(3)$ can be written as the telescoping sum

$$N(3) = [N(3) - N(2)] + N(2) = N^2(1) + 3.$$

As $\{N(2)=3, N(5)=10\} = \{N(2)=3, N^2(3)=7\}$, the conditional distribution of $N_2(1)$ is

$$N^2(1) \mid [\underbrace{N(2)=3}_{\text{get rid of this}}, N^2(3)=7] \quad \sim \quad N^2(1) \mid N^2(3)=7 \quad \sim \quad \text{Bin}(7, 1/3).$$

Finally, the conditional expected value of $N(3)$ is

$$\begin{aligned} \text{E}[N(3)|N(2)=3, N(5)=10] &= 3 + \text{E}[N^2(1)|N(2)=3, N^2(3)=7] \\ &= 3 + \frac{7}{3} \\ &= \boxed{5.3333}. \quad \textbf{(Answer: B)} \end{aligned}$$

□

Remark. In general, for $t_1 \le s \le t_2$ and a non-homogeneous Poisson process with mean value function $m(\cdot)$, the distribution of $N(s)$ conditional on $N(t_1)=A$ and $N(t_2)=B$ is

$$A + \text{Bin}\left(B - A, \frac{m(s)-m(t_1)}{m(t_2)-m(t_1)}\right).$$

The conditional expected value is

$$A + (B-A)\left[\frac{m(s)-m(t_1)}{m(t_2)-m(t_1)}\right] = \left[\frac{m(t_2)-m(s)}{m(t_2)-m(t_1)}\right]A + \left[\frac{m(s)-m(t_1)}{m(t_2)-m(t_1)}\right]B,$$

which is a weighted average of A and B.

1.2 Hazard Rate Function

KNOWLEDGE STATEMENTS

1c. Memoryless property of Exponential and Poisson

1d. Relationship between Exponential and Gamma

4a. Failure time random variables

4b. Cumulative distribution functions

4c. Survival functions

4d. Probability density functions

4e. Hazard functions and relationship to Exponential distribution

4f. Relationships between failure time random variables in the functions above

4g. Greedy algorithms

OPTIONAL SYLLABUS READING(S)

Ross, Section 5.2

In this section, we digress a bit to discuss a technical notion known as the failure rate function and some specialized results for the exponential distribution which will pave way for the further study of Poisson processes in the next section.

Failure rate function. We can look at the distribution of a random variable through its probability function and distribution function, as we usually do in our prior studies. A somewhat more colorful and sometimes more convenient way to describe the distribution of a random variable is furnished by the notion of failure rate function, which is defined as follows.

Given a continuous[ii] random variable X with distribution function F, its *failure rate function* $r(\cdot)$ (also known as *hazard rate function*) is defined as the ratio of its probability density function (p.d.f.) $f(\cdot)$ to its survival function $S(\cdot) = 1 - F(\cdot)$:

$$\boxed{r(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)}.} \tag{1.2.1}$$

An exam question expects you to calculate the failure rate function using (1.2.1) for a wide variety of distributions.

Case 1. If you are given the probability density function $f(\cdot)$, integrate it to obtain the survival function $S(\cdot)$.

[ii] Failure rate functions are traditionally defined only for continuous random variables.

Case 2. You may also be provided with the distribution function $F(\cdot)$. In this case, you can differentiate it to get back the probability density function. A quicker and alternative solution results from observing that

$$r(t) = \frac{f(t)}{S(t)} = \frac{-\frac{\mathrm{d}}{\mathrm{d}t}S(t)}{S(t)} = \boxed{-\frac{\mathrm{d}}{\mathrm{d}t}\ln S(t)} \tag{1.2.2}$$

because $\mathrm{d}\ln f(x)/\mathrm{d}x = f'(x)/f(x)$ for any positive function f with derivative f'.

These two cases are illustrated by the next two examples.

Example 1.2.1. (CAS Exam 3L Fall 2010 Question 1: Calculating the failure rate function given the density function) You are given the following density function:

$$f(t) = \frac{t^3}{c} \quad \text{for} \quad 0 \le t \le 10.$$

Calculate the failure rate function at $t = 5$.

A. Less than 0.035

B. At least 0.035, but less than 0.040

C. At least 0.040, but less than 0.045

D. At least 0.045, but less than 0.050

E. At least 0.050

Solution. The survival function is

$$S(t) = \int_t^{10} \frac{x^3}{c}\,\mathrm{d}x = \frac{10^4 - t^4}{4c}.$$

By (1.2.1),

$$r(5) = \frac{f(5)}{1 - F(5)} = \frac{5^3/c}{(10^4 - 5^4)/4c} = \frac{4}{75} = \boxed{0.0533}. \quad \textbf{(Answer: E)}$$

□

Remark. There is no need to determine the constant c, which will be canceled upon division.

Example 1.2.2. (CAS Exam 3L Spring 2008 Question 15: Calculating the failure rate function) You are given the following survival function:

$$S(t) = \mathrm{e}^{-5t^7}$$

Calculate $r(t)$, the failure rate function.

A. $5t^7$

B. $35t^6$

C. $35t^6\mathrm{e}^{-5t^7}$

D. $5t^7 \ln(35t^6)$

E. $\frac{35t^6\mathrm{e}^{-5t^7}}{1-\mathrm{e}^{-5t^7}}$

Solution 1. Using (1.2.2),

$$r(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\ln S(t) = -\frac{\mathrm{d}}{\mathrm{d}t}(-5t^7) = \boxed{35t^6}. \quad \textbf{(Answer: B)}$$

□

Solution 2. Alternatively, by differentiation, the probability density function is

$$f(t) = -\frac{\mathrm{d}}{\mathrm{d}t}S(t) = 35t^6\mathrm{e}^{-5t^7}, \quad t > 0.$$

By (1.2.1),

$$r(t) = \frac{f(t)}{S(t)} = \boxed{35t^6}. \quad \textbf{(Answer: B)}$$

□

[MINOR] How to make sense of the failure rate function? For a small $\mathrm{d}t$, we have the following approximate relationship:

$$r(t)\,\mathrm{d}t = \frac{f(t)\,\mathrm{d}t}{S(t)} \approx \frac{\Pr(t < X \leq t + \mathrm{d}t)}{S(t)} = \Pr(X \leq t + \mathrm{d}t \mid X > t).$$

Therefore, for a very small $\mathrm{d}t$, $r(t)\,\mathrm{d}t$ can be interpreted *loosely* as the probability that the random variable X, thought of the failure time of some entity, before "age" $t + \mathrm{d}t$ given that he/she/it survives "age" t. In this regard, $r(t)$ is a measurement of the *instantaneous* rate of failure. However, bear in mind that:

> *Failure rate function itself is never a probability! As such, $r(\cdot)$ is a conditional probability density function.*

As a simple example, consider a man aged 50 whose lifetime random variable has a failure rate function equal to 0.0044 per year at $t = 50$. By setting $\mathrm{d}t$ to be one day, or $1/365 = 0.002740$ year, we deduce that the *approximate* probability that the man dies on his 50th birthday *(poor guy!)* is $0.0044(0.002740) = 1.2 \times 10^{-5}$.

Relationship between failure rate function and other distributional quantities. Using (1.2.1) or (1.2.2), one can obtain the failure rate function of a certain continuous distribution, as we did in Examples 1.2.1 and 1.2.2. Conversely, we can also retrieve the underlying distribution function or survival function from a given failure rate function. The precise formula is given by

$$\boxed{S(t) = \exp\left(-\int_0^t r(s)\,\mathrm{d}s\right).} \tag{1.2.3}$$

Combining (1.2.1) (or (1.2.2)) and (1.2.3), we conclude that there is a one-to-one correspondence between the distribution function or survival function of a random variable and its failure rate function, or equivalently, the failure rate function uniquely determines the distribution of a random variable.

IMPORTANT EXAM ITEM

An exam question expects you to go between the failure rate function, survival function and probability density function efficiently using (1.2.1), (1.2.2) and (1.2.3).

Example 1.2.3. (Given $r(t)$, find...) The lifetime of a particular type of a newly purchased electronic product is modeled by a linear failure rate function given by

$$r(t) = \frac{1}{9}(t+2), \quad t \geq 0.$$

Calculate the 75th percentile of the lifetime of a new electronic product.

A. Less than 1.5

B. At least 1.5, but less than 2.0

C. At least 2.0, but less than 2.5

D. At least 2.5, but less than 3.0

E. At least 3.0

Solution. By (1.2.3), the survival function of a new electronic product is

$$\begin{aligned}
S(t) &= \exp\left(-\int_0^t r(s)\,\mathrm{d}s\right) \\
&= \exp\left\{-\frac{1}{9}\left[\frac{(s+2)^2}{2}\right]_0^t\right\} \\
&= \exp\left\{-\frac{1}{18}\left[(t+2)^2-2^2\right]\right\} \\
&= \exp\left[\frac{2}{9}-\frac{1}{18}(t+2)^2\right] \quad \text{for } t \geq 0.
\end{aligned}$$

The 75th percentile of the lifetime, denoted by t^*, satisfies $S(t^*) = 1 - 0.75 = 0.25$, so we set

$$\exp\left[\frac{2}{9}-\frac{1}{18}(t+2)^2\right] = 0.25 \quad \Rightarrow \quad t^* = \boxed{3.3808}. \quad \textbf{(Answer: E)}$$

□

Remark. Setting incorrectly $S(t^*) = 0.75$ would result in $t^* = 1.0296$, corresponding to Answer A.

[IMPORTANT!] Specialized properties of the exponential distribution. The second focus of this section is to develop some specialized properties of the exponential distribution, some of which will be useful to the further study of Poisson processes in the next section. Because the exponential distribution is a key character in the remainder of this section and arises in many other parts of Exam MAS-I as well, it pays to recall its simple definition.

RECALL

The probability density function of an exponential random variable X can be presented in two ways:

1. *(Rate parameterization)* Using λ as the *rate parameter*, we may write the density function as

$$f_X(x) = \lambda \mathrm{e}^{-\lambda x}, \quad x \geq 0.$$

The mean and variance of X are respectively

$$\mathrm{E}[X] = \frac{1}{\lambda} \quad \text{and} \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}.$$

Note that the rate of an exponential distribution is simply the reciprocal of its mean.

Such a rate parameterization of exponential distribution is usually handy for Poisson process calculations.

2. *(Scale parametrization)* If θ is the mean of the exponential distribution, then the probability density function becomes

$$f_X(x) = \frac{1}{\theta}\mathrm{e}^{-x/\theta}, \quad x \geq 0.$$

The mean and variance of X are respectively

$$\mathrm{E}[X] = \theta \quad \text{and} \quad \mathrm{Var}(X) = \theta^2.$$

This is the parameterization you can find in the table of distributions given in the exam.

It is important that you do not mix these two parameterizations up.

Example 1.2.4. (CAS Exam MAS-I Spring 2018 Question 7: Comparing two hazard rates) You are given the following information about an insurer:

- The amount of each loss is an exponential random variable with mean 2000.
- Currently, there is no deductible and the insurance company pays for the full amount of each loss.
- The insurance company wishes to introduce a deductible amount, d, to reduce the probability of having to pay anything out on a claim by 75%. The insurance company only pays the amount per loss exceeding the deductible.

- The insurance company assumes the underlying loss distribution is unchanged after the introduction of the deductible.

Calculate the minimum amount of deductible, d, that will meet the requirement of having 75% fewer claims excess of deductible.

A. Less than 2,000

B. At least 2,000, but less than 2,300

C. At least 2,300, but less than 2,600

D. At least 2,600, but less than 2,900

E. At least 2,900

Solution. The insurance company intends to set d so that the probability of having to pay a claim is $1 - 75\% = 0.25$. This requires $\mathrm{e}^{-d/2000} = 0.25$, or $d = \boxed{2772.5887}$. **(Answer: D)** □

Property 1. Exponential distribution is equivalent to a constant failure rate function: The exponential distribution is the only continuous distribution having a constant failure rate function.

$\Rightarrow$ If X is exponentially distributed with a rate of λ, then by (1.2.1)

$$r(t) = \frac{f(t)}{S(t)} = \frac{\lambda \mathrm{e}^{-\lambda t}}{\mathrm{e}^{-\lambda t}} = \lambda \quad \text{for all } t \geq 0.$$

$\Leftarrow$ Conversely, if the failure rate function $r(t)$ is constant at λ, then by (1.2.3),

$$S(t) = \exp\left(-\int_0^t r(s)\,\mathrm{d}s\right) = \mathrm{e}^{-\lambda t},$$

which is the survival function of an exponential distribution with a rate of λ.

Example 1.2.5. (CAS Exam MAS-I Spring 2018 Question 5: Comparing two hazard rates) You are given:

- Computer lifetimes are independent and exponentially distributed with a mean of 24 months.
- Computer I has been functioning properly for 36 months.

- Computer II is a brand new and functioning computer.

Calculate the absolute difference between Computer I's failure rate and Computer II's failure rate.

A. Less than 0.01
B. At least 0.01, but less than 0.02
C. At least 0.02, but less than 0.03
D. At least 0.03, but less than 0.04
E. At least 0.04

Solution. The failure rate is the same no matter whether the computer is brand new or an old one. Therefore, the absolute difference between Computer I's failure rate and Computer II's failure rate is zero. **(Answer: A)** □

Property 2. **[Important!]** *Memoryless property:* Another way to express Property 1 is the *memoryless property* (also known as lack of memory property), which says that conditional on $X > t$ for any $t \geq 0$, the *translated* random variable $X - t$ follows the same distribution as the original random variable X unconditionally. Symbolically, we can write

$$\boxed{X - t \mid X > t \overset{d}{=} X.}$$

With X regarded as the lifetime of a certain device, the memoryless property implies that the propensity for the device to fail from time t onward given that the device has survived time t is the same as that of a brand-new device – the original device does not "remember" that it has lived t units. In particular, for any $x \geq 0$ and $t \geq 0$, we have

$$\Pr(X > x + t | X > t) = \Pr(\boxed{X - t} > x | X > t) = \Pr(\boxed{X} > x),$$

where in the second equality we replace the translated random variable $X - t$ by the original random variable X and get rid of the conditioning event $\{X > t\}$.

Example 1.2.6. (CAS Exam S Fall 2016 Question 3: Standard application of memoryless property) The time X to wait in line is an exponentially distributed random variable with mean 5 minutes.

Calculate the probability that the total waiting time will be longer than 30 minutes from the time that individual arrived in line, given that the wait has already been 20 minutes.

A. Less than 0.1

B. At least 0.1, but less than 0.2

C. At least 0.2, but less than 0.3

D. At least 0.3, but less than 0.4

E. At least 0.4

Solution. By the memoryless property,

$$\begin{aligned}\Pr(X > 30|X > 20) &= \Pr(X - 20 > 10|X > 20)\\ &= \Pr(X > 10)\\ &= \mathrm{e}^{-10/5} = \mathrm{e}^{-2} = \boxed{0.1353}. \quad \textbf{(Answer: B)}\end{aligned}$$

□

Example 1.2.7. (CAS Exam S Fall 2015 Question 5: Exponential loss in the presence of a deductible) You are given the following information:

- The amount of damage involved in a home theft loss is an Exponential random variable with mean 2,000.
- The insurance company only pays the amount exceeding the deductible amount of 500.
- The insurance company is considering changing the deductible to 1,000.

Calculate the absolute value of the change in the expected value of the amount the insurance company pays per theft loss by changing the deductible from 500 to 1,000.

A. Less than 330

B. At least 330, but less than 350

C. At least 350, but less than 370

D. At least 370, but less than 390

E. At least 390

Ambrose's comments: This past exam problem was motivated from Example 5.4 of Ross.

Solution.
- For the deductible of 500, the expected value of the payment is $\mathrm{E}[(X - 500)_+]$, where $(\cdot)_+$ is the positive part function defined by $x_+ =$

$\max(x, 0)$ for any real x. By the law of total probability,

$$\begin{aligned}
\mathrm{E}[(X-500)_+] &= \mathrm{E}[\underbrace{(X-500)_+}_{0}|X<500]\,\Pr(X<500) \\
&\quad + \mathrm{E}[\underbrace{(X-500)_+}_{X-500}|X>500]\,\Pr(X>500) \\
&= \mathrm{E}[X-500|X>500]\,\Pr(X>500).
\end{aligned}$$

Next, the memoryless property of the exponential distribution says that conditional on X being greater than 500, $X - 500$ has the same distribution as the original exponential distribution with a mean of $2,000$. Therefore, $\mathrm{E}[X - 500|X > 500] = \mathrm{E}[X] = 2,000$, and

$$\mathrm{E}[(X-500)_+] = 2,000\mathrm{e}^{-500/2,000} = 1,557.6016.$$

- If the deductible is raised to 1,000, then likewise the expected value of the payment is

$$\mathrm{E}[(X-1,000)_+] = 2,000\mathrm{e}^{-1,000/2,000} = 1,213.0613.$$

The absolute value of the difference between the two expected values is $|1,557.6016 - 1,213.0613| = \boxed{344.5403}$. **(Answer: B)** □

Remark. You can also work out the problem by using the formula $\mathrm{E}[(X-d)_+] = \int_d^\infty S_X(x)\,\mathrm{d}x$, which is true for any random variable X with survival function S_X (not necessarily exponential).

The memoryless property admits some amazing and lesser-known extensions:

(a) It is possible to replace the reference point t by another independent non-negative random variable Y (not necessarily exponential), so that

$$X - Y \mid X > Y \overset{d}{=} X,$$

the right-hand side of which does not depend on Y. As a result, the conditional random variable $X - Y|X > Y$ is surprisingly independent of Y. An interesting by-product of this fact is that if X_1 and X_2 are independent exponential random variables, then

$$\boxed{\min(X_1, X_2) \quad \text{and} \quad \max(X_1, X_2) - \min(X_1, X_2)}$$

are independent[iii] random variables. This result is useful even in other parts of

[iii] Of course, $\max(X_1, X_2)$ and $\min(X_1, X_2)$ themselves are not independent, as $\min(X_1, X_2) \le \max(X_1, X_2)$ always.

Exam MAS-I.

(b) The memoryless property can even be extended to multi-dimensions involving several independent exponential random variables. Specifically, let X_1 and X_2 be two exponential random variables, and Y is another non-negative (not necessarily exponential) random variable. If X_1, X_2 and Y are mutually independent, then the *pair* of conditional translated random variables (with Y acting as the reference point) has the same distribution as the original unconditional pair:

$$\underbrace{(X_1 - Y, X_2 - Y)}_{\text{translated by } Y} \mid (X_1 > Y, X_2 > Y) \stackrel{d}{=} (X_1, X_2).$$

As a consequence, for $u \geq 0$ and $v \geq 0$,

$$\begin{aligned}\Pr(\boxed{X_1 - Y} > u, \boxed{X_2 - Y} > v \mid X_1 > Y, X_2 > Y) &= \Pr(\boxed{X_1} > u, \boxed{X_2} > v) \\ &= \Pr(X_1 > u)\Pr(X_2 > v).\end{aligned}$$

These generalized memoryless properties make some complex probabilities and expectations involving a group of independent exponential random variables easy without the need for explicit integrations.

Example 1.2.8. (The use of bivariate memoryless property) Let X_1, X_2 and X_3 be independent exponential random variables with parameters $\theta = 10$, $\theta = 20$ and $\theta = 30$ respectively.

It is known that X_3 is the smallest among X_1, X_2 and X_3.

Calculate the variance of $X_1 + X_2$.

A. Less than 350
B. At least 350, but less than 400
C. At least 400, but less than 450
D. At least 450, but less than 500
E. At least 500

Solution. Using the bivariate memoryless property presented above, we have

$$\begin{aligned} & \mathrm{Var}(X_1 + X_2 | X_3 = \min(X_1, X_2, X_3)) \\ = & \mathrm{Var}(\boxed{X_1 - X_3} + \boxed{X_2 - X_3} | X_1 > X_3, X_2 > X_3) \\ = & \mathrm{Var}(\boxed{X_1} + \boxed{X_2}).\end{aligned}$$

As X_1 and X_2 are independent, we further have

$$\begin{aligned}\mathrm{Var}(X_1 + X_2 | X_3 = \min(X_1, X_2, X_3)) &= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) \\ &= \theta_1^2 + \theta_2^2 \\ &= \boxed{500}. \quad \textbf{(Answer: E)}\end{aligned}$$

□

Property 3. Sum of i.i.d. exponential random variables gives rise to a gamma random variable: If $X_1, X_2, \ldots, X_n$ are independent exponential random variables with a common rate of λ[iv], then their sum $S = \sum_{i=1}^n X_i$ is a $\boxed{\text{gamma}}$ random variable with parameters $\alpha = n$ and $\theta = 1/\lambda$ (see the table of distributions). This fact is important for the study of event times of a Poisson process in Section 1.3.

Property 4. Minimum of independent exponential random variables is also exponential: If $X_1, X_2, \ldots, X_n$ are independent exponential random variables with respective rates $\lambda_1, \lambda_2, \ldots, \lambda_n$, then $\boxed{\min(X_1, X_2, \ldots, X_n)}$ also follows an exponential distribution, with a rate of $\boxed{\sum_{i=1}^n \lambda_i}$. This can be easily shown, for any $t \geq 0$, by

$$\begin{aligned}
\Pr(\min(X_1, X_2, \ldots, X_n) > t) &= \Pr(X_1 > t, X_2 > t, \ldots, X_n > t) \\
&\overset{\text{(independence)}}{=} \Pr(X_1 > t)\Pr(X_2 > t)\cdots\Pr(X_n > t) \\
&= e^{-\lambda_1 t} \times e^{-\lambda_2 t} \times \cdots \times e^{-\lambda_n t} \\
&= e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_n)t},
\end{aligned}$$

which is the survival function of an exponential distribution with a rate of $\sum_{i=1}^n \lambda_i$.

Example 1.2.9. (SAD STORY! A minimum of exponential random variables in disguise) Donald and Daisy are two ducks born on the same day. They love each other so much that if one dies, the other will drink a deadly poison immediately to die too.

The natural lifetimes of ducks are independent exponential random variables with mean 2 years.

If Donald and Daisy are to live for a further period of T years until they die together, calculate the expected value of T.

A. 0.5

B. 1

C. 2

D. 4

E. The answer is not given by A, B, C, or D.

Solution. Denote the natural lifetimes of Donald and Daisy by X_1 and X_2 respectively, both of which independently follow the exponential distribution with mean 2 years, or rate 1/2 per year. As soon as one of Donald and Daisy dies,

[iv] If the exponential rates are different, then it turns out that the distribution of $S = \sum_{i=1}^n X_i$ is a mixture of gamma distributions.

the other will also die together, so we have $T = \min(X_1, X_2)$, which is an exponential random variable with rate $1/2 + 1/2 = 1$. Therefore, $\mathrm{E}[T] = \boxed{1}$. **(Answer: B)** □

A prominent, but somewhat surprising fact concerning the minimum of independent exponential random variables is that $\min_{1\le i\le n} X_i$ and the *rank ordering* of the X_i's are independent. This is a consequence of the memoryless property: For any ordering $X_{i_1} < X_{i_2} < \cdots < X_{i_n}$ for the X_i's and $t > 0$, we have

$$\begin{aligned} & \Pr\left(X_{i_1} < X_{i_2} < \cdots < X_{i_n} \,\middle|\, \min_{1\le i\le n} X_i > t\right) \\ = \quad & \Pr\left(X_{i_1} - t < X_{i_2} - t < \cdots < X_{i_n} - t \mid X_i > t \text{ for all } i\right) \\ \overset{\text{(memoryless)}}{=} \quad & \Pr(X_{i_1} < X_{i_2} < \cdots < X_{i_n}), \end{aligned}$$

where the last probability is free of t.

Example 1.2.10. [HARDER!] (Expectations conditional on ordering) Let X_1 and X_2 be independent exponential random variables with respective rates λ_1 and λ_2.

(a) Determine an expression for $\mathrm{E}[X_1|X_1 < X_2]$.

A. $1/\lambda_1$
B. $1/\lambda_1 + 1/\lambda_2$
C. $1/(\lambda_1 + \lambda_2)$
D. $1/\lambda_1 + 1/(\lambda_1 + \lambda_2)$
E. $1/\lambda_2 + 1/(\lambda_1 + \lambda_2)$

Solution. Conditional on $X_1 < X_2$, note that X_1 equals the minimum of X_1 and X_2, i.e.,

$$\mathrm{E}[X_1|X_1 < X_2] = \mathrm{E}[\min(X_1, X_2)|X_1 < X_2].$$

As $\min(X_1, X_2)$ is independent of the ordering of X_1 and X_2, the preceding conditional expectation is simply the unconditional expectation:

$$\mathrm{E}[\min(X_1, X_2)|X_1 < X_2] = \mathrm{E}[\min(X_1, X_2)].$$

Finally, as $\min(X_1, X_2)$ is also an exponential random variable with rate $\lambda_1 + \lambda_2$, we have

$$\mathrm{E}[\min(X_1, X_2)] = \boxed{\frac{1}{\lambda_1 + \lambda_2}}. \qquad \textbf{(Answer: C)}$$

□

Remark. Conditional on $X_1 < X_2$, X_1 has the same distribution as $\min(X_1, X_2)$ unconditionally.

(b) Determine an expression for $\mathrm{E}[X_2|X_1 < X_2]$.

A. $1/\lambda_1$
B. $1/\lambda_1 + 1/\lambda_2$
C. $1/(\lambda_1 + \lambda_2)$
D. $1/\lambda_1 + 1/(\lambda_1 + \lambda_2)$
E. $1/\lambda_2 + 1/(\lambda_1 + \lambda_2)$

Solution. Rewriting X_2 as the telescoping sum $X_2 = X_1 + (X_2 - X_1)$, we have

$$\begin{aligned}\mathrm{E}[X_2|X_1 < X_2] &= \mathrm{E}[X_1 + (X_2 - X_1)|X_1 < X_2] \\ &= \mathrm{E}[X_1|X_1 < X_2] + \mathrm{E}[X_2 - X_1|X_1 < X_2].\end{aligned}$$

The first conditional expectation has been evaluated in part (a), while the second one can be dealt with easily by the memoryless property using X_1 as the reference point:

$$\mathrm{E}[X_2 - X_1|X_1 < X_2] = \mathrm{E}[X_2] = \frac{1}{\lambda_2}.$$

In conclusion,

$$\mathrm{E}[X_2|X_1 < X_2] = \boxed{\frac{1}{\lambda_1 + \lambda_2} + \frac{1}{\lambda_2}}. \quad \textbf{(Answer: E)}$$

□

Remark. Note that $\max(X_1, X_2)$ is *not* independent of the ordering of X_1 and X_2, so $\mathrm{E}[X_2|X_1 < X_2] = \mathrm{E}[\max(X_1, X_2)|X_1 < X_2] = \mathrm{E}[\max(X_1, X_2)]$ is not true.

Example 1.2.11. [HARDER!] (CAS Exam S Spring 2017 Question 6: Conditional expected value) X and Y are two independent exponential random variables with hazard rates $\lambda_X = 2$ and $\lambda_Y = 8$, respectively.

Calculate the expected value of X, conditional on $1 < X < Y$.

A. Less than 1.20
B. At least 1.20, but less than 1.40
C. At least 1.40, but less than 1.60

D. At least 1.60, but less than 1.80

E. At least 1.80

Solution. We are asked to find

$$\mathrm{E}[X|1 < X < Y] = 1 + \mathrm{E}[X - 1|1 < X, 1 < Y, X - 1 < Y - 1].$$

The bivariate memoryless property says that

$$(X - 1, Y - 1)|(X > 1, Y > 1) \stackrel{d}{=} (X, Y).$$

Thus replacing each "$X - 1$" by "X" and "$Y - 1$" by "Y" while getting rid of the conditioning events $\{1 < X\}$ and $\{1 < Y\}$ yields

$$\begin{aligned}
\mathrm{E}\left[\boxed{X-1}\,|1 < X, 1 < Y, \boxed{X-1} < \boxed{Y-1}\right] &= \mathrm{E}[X|X < Y] \\
&\overset{\text{(Example 1.2.10 (a))}}{=} \frac{1}{2+8} \\
&= 0.1.
\end{aligned}$$

The answer is $1 + 0.1 = \boxed{1.1}$. **(Answer: A)** □

Remark. To find the conditional variance of X, see Problem 1.2.15 on page 60.

Property 5. **[Important!]** *"Exponential race" probability*[v] *– Probability that one exponential is less than another independent exponential random variable:* If X_1 and X_2 are independent exponential random variables with respective rates λ_1 and λ_2, then

$$\boxed{\Pr(X_1 < X_2) = \frac{\text{rate of } X_1}{\text{rate of } X_1 + \text{rate of } X_2} = \frac{\lambda_1}{\lambda_1 + \lambda_2},}$$

which can be easily remembered by noting that:

- The numerator is the rate of the *dominated* exponential random variable X_1.
- The denominator is the sum of the two exponential rates.

More generally, if $X_1, X_2, \ldots, X_n$ are independent exponential random variables with respective rate $\lambda_1, \lambda_2, \ldots, \lambda_n$, then the probability that X_i is the smallest among the n

[v]This term can be found on page 96 of *Essentials of Stochastic Processes* (third edition), by Richard Durrett. This is not a required text for Exam MAS-I.

random variables, for $i = 1, 2, \ldots, n$, is

$$\begin{aligned}
\Pr\left(X_i = \min(X_1, X_2, \ldots, X_n)\right) &= \Pr\left(X_i < \underbrace{\min(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)}_{\text{exponential with rate } \sum_{j \neq i} \lambda_j \text{ (by Property 4)}}\right)\\
&= \frac{\lambda_i}{\lambda_i + \sum_{j \neq i} \lambda_j}\\
&= \boxed{\frac{\text{rate of } X_i}{\text{total rate}} = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j}}.
\end{aligned}$$

Example 1.2.12. (CAS Exam S Fall 2016 Question 7: Two groups of non-i.i.d. exponential random variables) You are given the following information about a watch with 6 different parts:

- There are 3 red wires with expected lifetimes of 50, 75, and 100.
- There are 3 yellow wires with expected lifetimes of 25, 50, and 75.
- The lifetimes of all wires are independent and exponentially distributed.

Calculate the probability that a red wire will break down before a yellow wire.

A. Less than 0.20

B. At least 0.20, but less than 0.25

C. At least 0.25, but less than 0.30

D. At least 0.30, but less than 0.35

E. At least 0.35

Ambrose's comments: This nice exam problem nicely combines Properties 4 and 5 in one single question.

Solution. The probability is equivalent to the probability that the minimum of the lifetimes of the 3 red wires is less than the minimum of the lifetimes of the 3 yellow wires. By Property 4:

- The minimum of the lifetimes of the 3 red wires is exponentially distributed with a rate of $1/50 + 1/75 + 1/100 = 13/300$.
- The minimum of the lifetimes of the 3 yellow wires is exponentially distributed with a rate of $1/25 + 1/50 + 1/75 = 11/150$.

Because the two minimums are independent exponential random variables, it follows from the "exponential race probability" that the required probability is

$$\frac{\overbrace{13/300}^{\text{red wire fails first}}}{13/300+11/150}=\frac{13}{35}=\boxed{0.3714}. \quad \textbf{(Answer: E)}$$

□

Example 1.2.13. (CAS Exam S Fall 2016 Question 5: Expected completion time) A call center currently has 2 representatives and 2 interns who can handle customer calls. If all representatives including interns are currently on a call, an incoming call will be placed on hold until a representative or intern is available.

You are given the following information:

- For each representative, the time taken to handle each call is given by an exponential distribution with a mean value equal to 1.
- For each intern, the time taken to handle each call is given by an exponential distribution with mean 2.
- Handle times are independent

A customer calls the call center and is placed on hold, and is the first person in line.

Calculate the expected time to complete the call (including both hold time and service).

A. Less than 1.2
B. At least 1.2, but less than 1.4
C. At least 1.4, but less than 1.6
D. At least 1.6, but less than 1.8
E. At least 1.8

Solution. As hinted at the end of the question, the time to complete the call consists of the hold time and the service time.

- The hold time, as the minimum of four independent exponential random variables, is exponential with rate $\underbrace{1/1+1/1}_{\text{representatives}}+\underbrace{1/2+1/2}_{\text{interns}}=3$ and therefore with mean $1/3$.

- The service time depends on whether the call is handled by a representative or an intern. By the same argument in Example 1.2.12, the probability that it is a representative is 2/3 and the probability that it is an intern is 1/3. By the law of total probability, the expected service time is

$$\frac{2}{3} \times \underbrace{1}_{\text{representative}} + \frac{1}{3} \times \underbrace{2}_{\text{intern}} = \frac{4}{3}.$$

The expected total time is $1/3 + 4/3 = \boxed{5/3 = 1.6667}$. **(Answer: D)** □

Remark. (i) A highly related problem with three servers and general exponential rates μ_1, μ_2, μ_3 is Exercise 5.25 of Ross, or Problem 1.2.27 on page 71.

(ii) If there are n servers with rates $\lambda_1, \ldots, \lambda_n$ to handle your call, then the expected time to complete the call can be shown to be $\frac{n+1}{\lambda_1+\lambda_2+\cdots+\lambda_n}$. You can verify your answer with $n = 4$ and $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = \lambda_4 = 0.5$.

(iii) Ross has a number of interesting and highly nontrivial problems of this sort. Some representative ones are solved in the end-of-section problems.

The "exponential race" probability can be combined with the memoryless property to determine some interesting and otherwise intractable probabilities expeditiously. Here is a typical example.

Example 1.2.14. (One exponential r.v bigger than the sum of other two) You are given that X_1, X_2, and X_3 are independent exponential random variables with hazard rates $\lambda_1 = 2$, $\lambda_2 = 4$, and $\lambda_3 = 8$, respectively.

Calculate $\Pr(X_3 > X_1 + X_2)$.

A. Less than 0.10

B. At least 0.10, but less than 0.15

C. At least 0.15, but less than 0.20

D. At least 0.20, but less than 0.25

E. At least 0.25

Solution. In terms of a conditional probability, our target of interest is

$$\begin{aligned}
\Pr(X_3 > X_1 + X_2) &= \Pr(\overbrace{X_3 > X_1 + X_2}^{\text{this implies } X_3 > X_1}, X_3 > X_1) \\
&= \Pr(X_3 > X_1 + X_2 | X_3 > X_1) \Pr(X_3 > X_1) \\
&= \Pr(X_3 - X_1 > X_2 | X_3 > X_1) \Pr(X_3 > X_1) \\
&\overset{\text{(memoryless)}}{=} \Pr(X_3 > X_2) \Pr(X_3 > X_1) \\
&\overset{\text{(exp. race)}}{=} \frac{\lambda_2}{\lambda_2 + \lambda_3} \times \frac{\lambda_1}{\lambda_1 + \lambda_3},
\end{aligned}$$

where the probabilistic identity $\Pr(A \cap B) = \Pr(B|A)\Pr(A)$ is used in the second equality. At $\lambda_1 = 2$, $\lambda_2 = 4$, and $\lambda_3 = 8$,

$$\Pr(X_3 > X_1 + X_2) = \frac{4}{4+8} \times \frac{2}{2+8} = \frac{1}{15} = \boxed{0.0667}. \quad \textbf{(Answer: A)}$$

□

Remark. (For integration geeks only) You can also compute the probability by triple integration using the joint probability density function of (X_1, X_2, X_3), but this requires much more tedious work.

[HARDER!] Application – Greedy algorithms. The syllabus covers Greedy Algorithms, which involve applications of the above properties of exponential distribution. Although the syllabus places Greedy Algorithms in the section of inter-arrival times of Poisson processes (Section 1.3 of this study manual), they are covered here for coherence. The details of Greedy Algorithms are a bit complicated, so you need to be patient.

Greedy Algorithms are designed to tackle the following problem. Suppose that n jobs are to be assigned to n people. A cost of C_{ij} is borne when person i is assigned to job j, for $i = 1, \ldots, n$ and $j = 1, \ldots, n$. Under the assumption that the C_{ij}'s form a set of n^2 independent exponential random variables with mean θ, the issue is how the n jobs should be assigned so as to minimize the *total expected cost*.

Two algorithms are proposed.

- *Greedy Algorithm A:* The first algorithm is simple and involves looking at each person one by one. It proceeds as follows.

  Step 1. We start by assigning to person 1 the job that leads to the smallest cost among $C_{11}, C_{12}, \ldots, C_{1n}$. Let's say that this is the j_1^{th} job. This results in bearing a cost of $C_{j_1} = \min_{1 \le j \le n} C_{1j}$. Then delete job j_1 – it has been allocated to person 1.

Step 2. Among the remaining $n-1$ jobs, assign the job, denoted by j_2, to person 2 resulting in the smallest cost among the C_{2j}'s for $j = 1, \ldots, n$, and $j \neq j_1$. Delete job j_2.

$\vdots$ $\vdots$

Step $n-1$. Among the remaining 2 jobs, assign the job with the smaller cost to person $n-1$. Delete this job.

Step n. Only one job and one person are left. They are then matched.

Mathematically, the total cost of assignment can be represented as

$$\min_{1\le j\le n} C_{1j} + \min_{1\le j\le n, j\neq j_1} C_{2j} + \cdots + \min_{1\le j\le n, j\neq j_1,\ldots,j_{n-1}} C_{nj},$$

and all the C_{ij}'s are mutually independent. By Property 4 of the exponential distribution, $\min_{1\le j\le n} C_{1j}$, as a minimum of n independent exponential random variables, has an exponential distribution with a rate of $n\theta^{-1}$. Similarly, $\min_{1\le j\le n, j\neq j_1} C_{2j}$ is an exponential random variable with a rate of $(n-1)\theta$. Consequently,

$$\begin{aligned}&\text{E[total cost of Greedy Algorithm A]}\\ &= \theta\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{2} + \frac{1}{1}\right) = \boxed{\theta\sum_{i=1}^{n}\frac{1}{i}}.\end{aligned} \tag{1.2.4}$$

Example 1.2.15. (CAS Exam S Spring 2016 Question 1: Expected total cost) Alice, Bob and Chris are hired by a firm to drive three vehicles - a bus, a taxi and a train (only one driver is needed for each). Each employee has different skills and requires different amounts of training on each vehicle. The cost to train each employee i for vehicle j, $C_{i,j}$, is an independent exponential random variable with mean of 200.

To minimize the total cost of training, the firm uses the following procedure to assign the employees to their vehicle:

- Alice is assigned to the vehicle which minimizes her training cost, $C_{Alice,j}$.
- Bob is then assigned one of the two remaining vehicles which minimizes $C_{Bob,j}$.
- Chris is then assigned the remaining vehicle.

Calculate the firm's expected total cost of training these three employees, if it uses this assignment algorithm.

A. Less than 200

B. At least 200, but less than 300

C. At least 300, but less than 400

D. At least 400, but less than 500

E. At least 500

Solution. By (1.2.4), the expected total cost of training is

$$200\left(\frac{1}{3}+\frac{1}{2}+\frac{1}{1}\right)=\frac{1,100}{3}=\boxed{366.67}. \quad \textbf{(Answer: C)}$$

□

- *Greedy Algorithm B:* The second algorithm is harder and adopts a more global perspective. At each step, it seeks to match the person and job that have the minimal cost among all unassigned people and jobs. It starts by considering all n^2 costs and selects the pair (i_1, j_1) such that C_{i_1,j_1} is the smallest among the n^2 costs. Next, eliminate (i_1, j_1) (i.e., delete person i_1 and job j_1) and choose the pair (i_2, j_2) such that C_{i_2,j_2} is the smallest among the remaining $(n-1)^2$ costs. The procedure is repeated until every person receives his/her job assignment.

  To analyze the expected total cost of Algorithm B, let C_i be the cost associated with the i^{th} person-job pair. As C_1 is the minimum of n^2 costs, it possesses an exponential distribution with a rate of $n^2\theta$ (by Property 4). Now you may be tempted to assert that C_2, being the minimum of $(n-1)^2$ costs, is exponentially distributed having a rate of $(n-1)^2\theta$. The truth is that the remaining $(n-1)^2$ costs, by definition, are all greater than $C_1 = C_{i_1,j_1}$. By the memoryless property, the amounts by which these $(n-1)^2$ costs exceed C_1 remain to be independent exponential random variables with mean θ, so these C_{ij}'s are distributed according to

$$C_{ij} \overset{d}{=} C_1 + \text{Exponential with rate } \theta \quad \text{for all } i \neq i_1, j \neq j_1.$$

  It follows that

$$C_2 = \min_{\substack{1\le i\le n, 1\le j\le n \\ i\neq i_1, j\neq j_1}} C_{ij} \overset{d}{=} C_1 + \text{Exponential with rate } (n-1)^2\theta,$$

  so the expected value of C_2 is related to that of C_1 via

$$\mathrm{E}[C_2] = \mathrm{E}[C_1] + \frac{\theta}{(n-1)^2} = \theta\left[\frac{1}{n^2}+\frac{1}{(n-1)^2}\right].$$

  Inductively, we have

$$\mathrm{E}[C_{i+1}] = \mathrm{E}[C_i] + \frac{\theta}{(n-i)^2} \quad \text{for } i = 1, \ldots, n-1,$$

or more explicitly,

$$
\begin{aligned}
\mathrm{E}[C_1] &= \frac{\theta}{n^2}, \\
\mathrm{E}[C_2] &= \theta\left[\frac{1}{n^2}+\frac{1}{(n-1)^2}\right], \\
\mathrm{E}[C_3] &= \theta\left[\frac{1}{n^2}+\frac{1}{(n-1)^2}+\frac{1}{(n-2)^2}\right], \\
&\vdots \\
\mathrm{E}[C_n] &= \theta\left[\frac{1}{n^2}+\frac{1}{(n-1)^2}+\frac{1}{(n-2)^2}+\cdots+\frac{1}{2^2}+\frac{1}{1^2}\right].
\end{aligned}
$$

The total expected cost, after adding $\mathrm{E}[C_1],\ldots,\mathrm{E}[C_n]$, is

$$\mathrm{E}[\text{total cost of Greedy Algorithm B}] = \theta\left[\frac{n}{n^2}+\frac{n-1}{(n-1)^2}+\cdots+\frac{1}{1^2}\right] = \boxed{\theta\sum_{i=1}^{n}\frac{1}{i}},$$

which is the same as that of Algorithm A.

In fact, after all the fuss, a stronger conclusion can be drawn:

The distributions of the total cost happen to be the same for Algorithm A and Algorithm B.

This is the content of Exercise 5.27 of Ross. Therefore, if an exam question is set on Greedy Algorithm B, you may legitimately "cheat" and assume that the simpler Greedy Algorithm A is used. In particular, the variance of the total cost is

$$\boxed{\mathrm{Var}(\text{total cost}) = \theta^2\sum_{i=1}^{n}\frac{1}{i^2}.}$$

However, it is important to understand the rationale that underlies Algorithm B, because an exam question may apply these considerations to a slightly different context, e.g., Exercise 5.17 of Ross, which is solved below.

Example 1.2.16. (Ross, Exercise 5.17: A variant of Greedy Algorithm) *Please refer to the book for the question statements.*

Ambrose's comments: This problem applies Greedy Algorithm B to a related but different context involving the construction of links.

Solution. Let C_i be the cost of the i^{th} link for $i = 1,\ldots,n-1$. Being the minimum of $\binom{n}{2}$ independent unit-rate exponential random variables,

$$\mathrm{E}[C_1] = \binom{n}{2}^{-1}.$$

By the memoryless property, the amounts by which the costs of other links exceed C_1 are independent unit-rate exponential random variables. Then, because the next link is constructed connecting any of the two cities connected by the first link and any of the remaining $(n-2)$ cities,

$$C_2 = C_1 + \underbrace{(C_2 - C_1)}_{\substack{\text{min. of } 2(n-2) \text{ unit-rate} \\ \text{exponential r.v.}}},$$

has an expected value of

$$\mathrm{E}[C_2] = \mathrm{E}[C_1] + \frac{1}{2(n-2)}.$$

More generally,

$$C_i = C_{i-1} + \underbrace{(C_i - C_{i-1})}_{\substack{\text{min. of } i(n-i) \text{ unit-rate} \\ \text{exponential r.v.}}} \quad \text{and } \mathrm{E}[C_i] = \mathrm{E}[C_{i-1}] + \frac{1}{i(n-i)}, \quad i = 2, \ldots, n-1.$$

(a) When $n = 3$, the expected cost is

$$\mathrm{E}[C_1] + \mathrm{E}[C_2] = 2\binom{3}{2}^{-1} + \frac{1}{2(3-2)} = \boxed{\frac{7}{6}}.$$

(b) When $n = 4$, the expected cost is

$$\mathrm{E}[C_1] + \mathrm{E}[C_2] + \mathrm{E}[C_3] = 3\binom{4}{2}^{-1} + \frac{1}{2(4-2)} + \frac{1}{3(4-3)} = \boxed{\frac{13}{12}}.$$

□

Remark. Unlike the Greedy Algorithms, the first cost C_1 in this example has a different rate structure.